Approches innovantes en RMN biomoléculaire: Cinétiques moléculaires par RMN rapide et paroi bactérienne par RMN du solide

Thomas Kern

UNIVERSITE JOSEPH FOURIER – GRENOBLE 1
ECOLE DOCTORALE DE PHYSIQUE

Thèse pour l'obtention du grade de
DOCTEUR DE L'UNIVERSITE JOSEPH FOURIER

Spécialité: Physique pour les Sciences du Vivant

présentée par

Thomas KERN

Approches innovantes en RMN biomoléculaire: Cinétiques moléculaires par RMN rapide et paroi bactérienne par RMN du solide

Tackling challenging systems by biomolecular NMR spectroscopy: Molecular kinetics by fast NMR and bacterial cell wall by solid-state NMR

Directeur de thèse: Dr. Jean-Pierre SIMORRE
Co-Directeur de thése: Dr. Bernhard BRUTSCHER

Thèse soutenue publiquement le 22 octobre 2009 devant le jury:

Dr. Bernhard BRUTSCHER	Co-Directeur de thèse
Dr. Erick DUFOURC	Rapporteur
Pr. Dr. Alain MILON	Rapporteur
Dr. Jean-Pierre SIMORRE	Directeur de thèse
Dr. Waldemar VOLLMER	Examinateur
Pr. Dr. Winfried WEISSENHORN	Examinateur

Gewidmet meinem Zwillingsbruder Matthias

REMERCIEMENTS

J'aimerais commencer par remercier Pr. Erick Dufourc et Pr. Alain Milon, d'avoir accepté d'être rapporteur de ma thèse et Dr. Waldmar Vollmer et Pr. Winfried Weissenhorn d'avoir été examinateur lors de ma soutenance de thèse.

Après trois années de thèse, je quitte le laboratoire en sachant que j'ai eu beaucoup de chance ces trois dernières années. J'ai eu beaucoup de chance de faire partie du laboratoire de RMN où je n'ai pas seulement beaucoup appris, mais où je me suis toujours senti à ma place. J'ai eu la chance d'avoir un travail que j'ai beaucoup aimé avec des gens que j'ai retrouvés tous les jours avec plaisir et qui sont devenus des amis. Sans eux cette thèse n'aurait pas été possible.

Tout d'abord j'aimerais remercier Jean-Pierre Simorre qui m'a fait venir au laboratoire et qui a accepté le grand défit de faire de moi un docteur. J'ai beaucoup apprécié de travailler avec lui et il a toujours été plus qu'un simple chef, mais un ami. Merci pour tout ce que tu m'as appris, pour avoir toujours été là quand j'en ai eu besoin et pour avoir été un chef "ami" avec qui il a toujours été rigolo de discuter, qui a supporté des nombreux discours sur l'escalade et qui a même pris le risque de se balader en montagne avec moi. J'ai eu un chef courageux!

Je tiens aussi à remercier Bernhard Brutscher pour m'avoir donné la possibilité de travailler avec lui sur des sujets plus axés sur la spectroscopie. Merci beaucoup pour toutes ces journées passées sur le spectromètre. Tout le monde n'a pas la chance d'avoir un "deuxième" chef d'une telle qualité scientifique et humaine. ça a été un grand plaisir!

Pendant cette thèse, on s'est lancé dans une aventure toute nouvelle pour notre groupe, la RMN du solide. C'est ici que je ne peux pas assez remercier Sabine Hediger et sa thésarde Mathilde Giffard nos experts de la RMN du solide qui nous ont accompagné tout le long du projet. Sans elles ce projet n'aurait jamais été possible.

Vu la diversité de mes projets il y avait de nombreux acteurs impliqués qui ont tous contribué à ce travail. Parmi eux il y a spécialement Waldemar Vollmer de la Medical School de Newcastle que je veux remercier. Il était là au tout début du projet sur la paroi bactérienne et y a contribué tout au long de ma thèse avec ses conseils et sa connaissance extraordinaires du domaine de la paroi bactérienne. Il m'a aussi invité à Newcastle où j'ai passé un été inoubliable à apprendre à purifier la paroi cellulaire. Merci à toi Waldemar, mais aussi à Patrick et Khai, pour l'aide à la purification des sacculi, à Robyn et Richard pour m'avoir aidé avec les mutants, et à Manuel, Alise, Biserka, Pati, Danny, Kristoffer, Henrik, Thomas, Mickel I, Mickel II, Rok, Marc, Katrin, Simon, Petra... pour avoir rendu mon passage là-bas si agréable!

Dans le contexte du projet sur la paroi bactérienne je tiens aussi à remercier Bernhard Joris et Ana Amoroso de l'université de Liège, qui ont été impliqués dans le projet d'interaction de la paroi bactérienne avec YajG et qui nous ont beaucoup aidé avec leur expérience et spécialement Ana pour son aide à la purification de la paroi cellulaire de nombreuses souches.

Merci à Paul Schanda pour m'avoir donné la possibilité de continuer à travailler sur son petit bébé la SOFAST (qui a dû récemment faire place à la encore plus belle petite Emilia, encore félicitations au papa et bien sûr à la maman). Merci aussi à Maayan Gal et Lucio Frydman pour le travail que l'on a fait ensemble dans le contexte de l'UltraSOFAST et spécialement à Maayan pour tout le bon temps que l'on a passé ensemble au labo et sur les pistes de ski lors de son passage à Grenoble.

Je remercie Catherine Bougault pour avoir été ma prof de RMN et pour avoir été ensuite toujours là pour moi et tous les étudiants tout court, pour toutes ces discussions, pour avoir répondu à mes nombreuses questions, pour avoir été impliqué dans tout ce que Jean-Pierre et moi avons fait et pour avoir toujours donné un avis critique sur notre travail.

Merci à Ewen Lescoop de m'avoir initié aux grands secrets du processing et pour m'avoir accompagné sur mes premiers enregistrements de spectres.

Je tiens également à remercier Isabelle Ayala qui fait que notre wetlab ne parte pas dans le chaos le plus total et qui a su m'empêcher à en faire un. Merci Isa pour avoir fait ce travail de titan pour nous en étant toujours si charmante et drôle. Un grand merci aussi à Cécile Giustini et Ombeline Pessey qui m'ont beaucoup aidé avec tout le travail à la paillasse.

Pour avoir été là au quotidien, pour avoir toujours été disponible pour des discussions scientifiques et autre, pour toutes ces petites choses qui ont rendu ma thèse si agréable je veux remercier tous les membres permanents du LRMN, Jérôme Boisbouvier, Dominique Marion, Pierre Gans, Beate Bersch, Eric Condamine. En particulier je veux remercier Adrien Favier pour tout ce qu'il fait pour que les spectromètres marchent et pour m'avoir aidé avec tous mes soucis informatiques.

Merci à notre groupe de thésard et postdoc, Paul, Michael, Maria-Rosa, Julien, Guillaume, Thomas, Lorianne, Carlos, Fito, Nico...Particulièrement à Klaartje pour nous avoir toujours motivé à bouger et pour cet appartement magnifique que j'ai eu les deux dernières années !

Merci aussi à nos voisins du FDP Loïc (qui a d'ailleurs fait en sort que ce que vous lisez reste compréhensible), Antoine, Valery. Particulièrement à Malene pour m'avoir remonté le moral chaque fois quelque chose ne marchait pas et pour tout ce que tu as fait pour moi dans ces dernières années. Et merci à Martin Blackledge pour toutes ces discussions scientifiques et surtout pour toutes ces discussions "grimpe" et ton topo d'Ombleze!

Un grand merci aussi à nos collaborateurs au sein de l'IBS Andrea, Anne-Marie, André, Soizic, Clothilde, Frank, Thierry... qui font qu'il n'y a rien qu'on ne peut pas faire chez nous.

Je remercie Vincent Forge chez qui j'ai passé mon stage de M1 et de M2. Il m'a initié à la recherche et m'a montré que faire de la bonne recherche et de rigoler en même temps n'est pas incompatible!

Bien sûr que je n'ai pas fait que travailler pendant ces trois dernières années et j'ai trouvé pendant ce temps sur Grenoble ma passion pour la montagne et surtout pour l'escalade. Je veux remercier Paul Schanda, Klaartje Houben, Nicolas Duraffourg, Adrien Favier, Jean-Pierre Simorre de m'y avoir initié et d'avoir accompagné mes premiers pas sur tous ces sommets. Et surtout Klaartje pour m'avoir amené la première fois à l'escalade ici, tu n'es peut-être même pas consciente de ce que tu as déclenché. Au fur et au mesure de ma découverte du milieu de l'escalade de Grenoble je suis tombé sur A.Blok et c'est devenu ma deuxième maison et la base de tous mes projets d'escalade. Merci Bullet, Hugo, Barns, Bruno, Cécile, Dams, Delphine, Fab, Francois, Jean-Jean, Jean-Marie, Ju, Joan, Marco, Matthieu, Maya, Mia, Michael, Nam, Nico, Rackam, Remy, Sylvain, Sylvie, Yann...

J’aimerais finir ces remerciements avec tous les gens qui font partie du projet européen EU-INTAFAR dans lequel le travail sur la paroi bactérienne c’est situé et par lequel il a été financé.

CONTENTS

"Ne vivez pour l'instant que vos questions. Peut-être, simplement en les vivant, finirez-vous un jour, par entrer insensiblement dans les réponses"

— Rainer-Maria Rilke

ABSTRACT

The thesis entitled "Tackling challenging systems by biomolecular NMR spectroscopy" is divided in two parts. The first part treats recent advances in the field of fast NMR. Standard multidimensional NMR methods required to yield sufficient atom-resolution are time-consuming rendering real-time analysis difficult. The SOFAST-HMQC experiment allows 2D spectra of macromolecules to be recorded in only a few seconds. This thesis is concerned with the optimization and extension of longitudinal relaxation enhanced pulse sequences in the context of SOFAST and BEST type experiments. SE-IPAP-SOFAST-HMQC is introduced, a new pulse sequence that provides comparable sensitivity to standard SOFAST-HMQC, while avoiding heteronuclear decoupling during ^{1}H detection, and thus significantly reducing the radio-frequency load of the probe during the experiment. The experiment is also attractive for fast and sensitive measurements of heteronuclear one-bond spin coupling constants. Further, a scheme for improved ultrafast data acquisition is presented yielding higher signal to noise and improved line shapes in single-scan 2D NMR spectra. In the context of longitudinal relaxation enhanced BEST experiments, it is shown that selective pulses can be replaced by simplified representations consisting of a combination of free evolution delays and ideal rf pulses. They reproduce well the effect of shaped pulses on a scalar-coupled heteronuclear spin system and make implementation of such pulse shapes in complex multipulse sequences straightforward. Finally we show a relaxation-dispersion experiment based on the BEST concept, illustrating that the relaxation enhancement concept by selective excitation can be extended to a wide range of NMR experiments.

The second part of this thesis deals with the bacterial cell wall of Gram-negative and Gram-positive bacteria. Because of its size and non cristalline character we applied solid state NMR to study it. Basic concepts of solid state NMR used in this context are introduced and applied to intact peptidoglycan sacculi of the Gram-negative bacterium *Escherichia coli* and the Gram-positive bacteria *B. subtilis*, *S. aureus*, and *S. pneumoniae*. The high-quality solid-state NMR spectra allow atom-resolved investigation of the peptidoglycan structure and dynamics. Flexibility and organization of the

peptidoglycan from the Gram-negative bacteria (*E. coli*) are compared with its counterpart from different Gram-positive bacteria. ^{13}C T_1 relaxation constants are used to analyze the impact of cross-linking, glycan chain length and peptidoglycan thickness on the flexbility of the peptidoglycan of different species. Determination of dynamical properties of the peptidoglycan can also be used to study peptidoglycan-protein interactions. This is shown on the example of the interaction of the lipoprotein YajG with the peptidoglycan of *E. coli*, studied by determining T_1 relaxation time constants of the peptidoglycan with and without the protein attached. Finally, the NMR concepts are extended to the cell wall of Gram-positive bacteria where the wall teichoic acids (WTAs) replace the function of the outer membrane of Gram-negative bacteria. ^{31}P solid-state NMR is applied to detect the phosphor groups of the teichoic acids. These data prove to be particularly well adapted to characterize and differentiate wall teichoic acids of different species. Further complexation with divalent ions is observed highlighting an important structural aspect of Gram-positive cell wall architecture.

RÉSUMÉ

La thèse intitulée "Approches innovantes en RMN biomoléculaire: Cinétiques moléculaires par RMN rapide et paroi bactérienne par RMN du solide" est divisée en deux parties. La première partie traite des progrès récents dans le domaine de la RMN à acquisition rapide. Les méthodes RMN multidimensionnelles requises pour l'obtention d'une résolution atomique élevée sont relativement couteuses en temps expérimental, ce qui complique considérablement leur application à l'analyse d'échantillons en temps réel. En revanche, l'expérience SOFAST-HMQC permet l'enregistrement de spectres 2D de macromolécules en seulement quelques secondes. Cette thèse porte sur l'optimisation et l'extension des séquences d'impulsions de type SOFAST et BEST à relaxation longitudinale augmentée. La nouvelle séquence d'impulsion SE-IPAP-SOFAST-HMQC est introduite, elle fournit une sensibilité comparable à SOFAST-HMQC tout en évitant le découplage hétéronucléaire pendant la détection proton; ceci permet de réduire sensiblement la charge de radio-fréquence sur la sonde pendant l'expérience. Cette séquence est également attrayante pour des mesures rapides et sensibles de constantes de couplage J hétéronucléaire. En outre, une séquence améliorée d'acquisition ultra-rapide des donnés est présentée fournissant une meilleure sensibilité et une forme de raie améliorée pour les spectres 2D enregistrés en une seule accumulation. Dans le contexte des expériences de relaxation longitudinale type BEST amélioré, il est démontré que les impulsions sélectives peuvent être remplacées par des représentations simplifiées composées de combinaisons de délais d'évolution libre et d'impulsions RF idéales. Elles reproduisent bien l'effet d'impulsions sélective sur un système de spins hétéronucléaires couplés et facilitent sensiblement leur implémentation dans des séquences multi-pulse complexes. Enfin, nous montrons une expérience de dispersion de relaxation basée sur le principe BEST, ce qui montre que le concept d'amélioration de relaxation par l'excitation sélective peut être étendu à un large éventail d'expériences de RMN.

La deuxième partie de cette thèse concerne la paroi cellulaire des bactéries Gram-négatives et Gram-positives. En raison de son poids moléculaire élevé et de son caractère non cristallin, nous avons appliqué la RMN du solide pour l'étudier. Les concepts de base de la RMN à l'état solide utilisés dans ce

contexte sont introduits et appliqués au peptidoglycane intact de la bactérie Gram-négative Escherichia coli ainsi qu'a celui des bactéries Gram positives *B. subtilis*, *S. aureus* et *S. pneumoniae*. La qualité exceptionnelle des spectres de RMN solide permettent l'étude, à résolution atomique, de la structure et de la dynamique du peptidoglycane. La flexibilité et l'organisation du peptidoglycane de *E. coli* sont comparées avec les caractéristiques du peptidoglycane de différentes bactéries Gram-positives. Les constantes de relaxation T_1 des spins ^{13}C sont utilisées pour analyser l'impact de la réticulation, de la longueur de chaîne et de l'épaisseur des glycanes sur la flexibilité du peptidoglycane de différentes espèces. La détermination des propriétés dynamiques du peptidoglycane peut aussi être utilisée pour étudier les interactions entre protéines et peptidoglycane. Ceci est illustré sur l'exemple de l'interaction entre la lipoprotéine YajG et le peptidoglycane d'*E. coli*, en comparant les constantes de temps de relaxation du peptidoglycane avec et sans la protéine. Enfin, les concepts présentés sont étendus à la paroi cellulaire des bactéries Gram-positives où les acides téchoïques qui se lient de manière covalente au peptidoglycan (WTA) remplacent la fonction de la membrane externe des bactéries à Gram négatif. La RMN du solide ^{31}P est appliquée à la détection des groupements phosphoriques des acides téchoïques. Ces données s'avèrent particulièrement bien adaptées à la caractérisation et différenciation des acides téchoïques d'espèces différentes. En outre la complexation avec des ions divalents est observée soulignant un aspect structurel important de l'architecture de la paroi cellulaire des bactéries Gram-positives .

Part I

INTRODUCTION TO FAST LIQUID STATE NMR

1

FAST LIQUID STATE NMR

1.1 INTRODUCTION

In structural biology, nuclear magnetic resonance (NMR) and crystallography are two major techniques used to study protein folds and structural modifications on ligand binding. While crystallography is a great tool for resolving atomic resolution protein structures of large complexes, it is limited to the study of frozen crystal structures. Although highly resolved structures are very useful they are only one aspect in the understanding of biological processes. Protein dynamics are at least as important as static structures and both aspects can be nicely addressed by nuclear magnetic resonance spectroscopy. Furthermore NMR provides atomic-level information without disturbing the chemical properties of the investigated object because it is only manipulating nuclear spins with electromagnetic fields. Consequently it has become an important tool in biological research, but also in a broad range of other fields as physical and material sciences, chemistry, and medicine (MRI).

In the last fifty years many powerful tools have been developed in the field of NMR to study molecular dynamics on the time scale of picoseconds to milliseconds. Examples of NMR observables providing dynamical information are T_1, T_2 relaxation times[64, 122], heteronuclear NOEs, relaxation-dispersion effects [51, 99, 100], residual dipolar couplings [11], and H/D exchange rates [19]. Therefore NMR has become a very powerful tool for studying dynamical properties of biological phenomena like protein binding, folding or structural exchange. While all these techniques are exploiting time averaged observable it gets more difficult when the aim of the study is to study slower processes. In this case real time NMR is required [147, 170, 183] and the time resolution depends on the experimental time required to record a single spectrum (figure 1a).

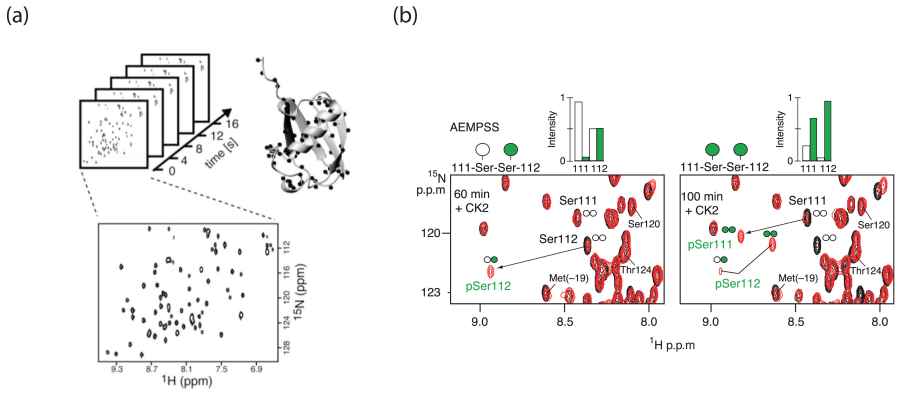


Figure 1 – (a) Example of how SOFAST HMQC spectra can be used to monitor in a real time mode time dependent spectral changes [147]. Each spectrum and each cross-peak monitor local structure and dynamics at different time points. (b) Phosphorylation of Ser112 and Ser111 of the peptide $XT_{107-132}GB1$ by the casein kinase 2 (CK2) monitored over time after 60 min (left) and 100 min (right) [155].

An example of this kind of study is the phosphorylation of two Serine residues (Ser112 and Ser111) in the peptide $XT_{107-132}GB1$ a substrate site within the viral SV40 large T antigen regulator by the casein kinase 2 (CK2). Wagner and his coworkers were using a repetition of ^{15}N-HSQC spectra to monitor this process over time in vitro and in vivo (*Xenopus oocytus*) (figure 1b) [155, 156]. They were able to show that the phosphorylation sites are processed in a stepwise manner by the kinase and that an intermediate substrate is released. Ser112 needs to be phosphorylated before Ser111 is modified, thus demonstrating that phosphorylation is not only dependent on the primary structure but also on its order of modification. This is only visible when distinct moments in the progress of modification can be observed and timely resolved. In the example of the experiments done in vitro single spectrum acquisition on a 50 µM ^{15}N-labeled substrate took about 43 min acquisition time on a 600 MHz spectrometer equipped with a cryo-probe. Spectra taken in vivo needed even 90 min of acquisition time because of the lower sensitivity due to the experimental conditions.

It speaks for itself that a time resolution of about 45 min is in the majority of cases not sufficient to monitor kinetics on the timescale of seconds to some minutes. Therefore a lot of effort has been invested to speed up NMR spectra acquisition [32]. The band-Selective Optimized-Flip-Angle Short-Transition (SOFAST) experiment [144] is an example that demonstrates how far one can go in accelerating spectra acquisition to Hertz rates (figure 1a).

The study of protein dynamics and structure would not be possible, without multidimensional NMR techniques which have become crucial for studying larger biological objects such as proteins [9]. Without pulse sequences developed to correlate several nuclei, structure definition of larger proteins would still be a difficult task. But adding an dimension also means a significantly increased number of scans needed. Adding one resonance to a 2D correlation spectrum using a reasonable resolution of 50 sample points in this dimension increases the time needed to acquire the spectrum by two orders of magnitude.

If additionally a good sensitivity is needed because of a higher protein size, low sample concentration, or needed buffer conditions (high salt concentration, high pH, detergents,...) this kind of experiments can therefore take hours to days. This problem of sensitivity and recording time has been addressed by the set of Band-selective Excitation Short Transient spectroscopy (BEST) experiments [94] allowing significantly faster recording of higher dimensional spectra and by the SOFAST sequence [144, 146, 148] for 2D correlation experiments.

The following part of this thesis will discuss recent concepts in the field of fast NMR techniques that benefit from the improved performances of NMR hardware in terms of sensitivity and resolution achieved by cryogenically cooled detection coils and preamplifiers, and high magnetic field strengths.

1.2 PRINCIPLES OF FAST NMR

All the information NMR can collect is contained in the quantum mechanical property of the studied molecule's atoms, called nuclear spin. It is collected by analyzing the response of those spins, placed in a static magnetic field, to radio-frequency electromagnetic waves. This response to spin excitation is generally represented as a spectrum showing the signal intensity as a function of frequency. Every signal raising in the spectrum can be described by its position, intensity, and line-shape. All the extractable information about various structural and dynamical properties of the object is based on these observables, thus the quality of the analysis is directly dependent on the precision and accuracy of their measurement. Precision in the case of NMR is mainly governed by the obtained *spectral resolution* and the *experimental sensitivity* of the experiment.

Spectral resolution

Resolution refers to the minimum difference between two individual signals that can be identified as individual resonances. Due to the narrow frequency range of the observable nuclei in NMR (^{1}H, ^{13}C, ^{15}N, and ^{31}P) one-dimensional NMR of large molecules suffers from severe resonance overlap. This makes it practicably impossible to extract any atomic-resolved information about the structure or dynamics of even small biomolecules. This problem has been solved by the introduction of multidimensional NMR increasing the resolution by correlating nuclear spins [4, 65]. Multidimensional NMR will be discussed in more detail later on in this section.

Experimental sensitivity

One other issue of NMR is its low experimental sensitivity due to the weak nature of the magnetic spin interactions measured in an experiment. It is defined as the ratio between the observed signal and the spectral noise normalized for time unity. In the last years sensitivity has progressively been enhanced by hardware developments such as higher magnetic fields and cryogenically cooled probes. Along with the hardware improvement sophisticated NMR pulse sequences were designed to minimize signal loss due to pulse imperfections, spin relaxation, solvent signals, thus maximizing the achievable signal.

Resolution and sensitivity necessary to extract the researched information determines therefore the minimum duration of an experiment. In other words exploiting the recent progress made in the field, and optimizing or developing new pulse sequences to gain resolution or sensibility also means to speed up data collection.

1.3 TIME REQUIREMENTS IN FOURIER TRANSFORM NMR

1.3.1 *Composition of one scan in a NMR experiment*

To find ways of accelerating data acquisition it is first of all useful to have a closer look at the parameters that are defining the length of a NMR experiment.

Figure 2 shows a very simplified representation of NMR data acquisition by Fourier transform NMR. One scan of an experiment can be separated in three main parts. In the first time period T_{rf}, of the pulse sequence, nuclear spins

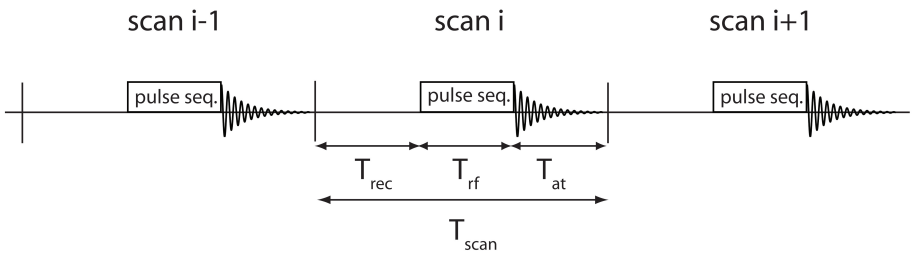


Figure 2 – Schematic representation of the time parameters of a NMR experiment. The overall time of an experiment can be reduced by either reducing the number of repetitions (scans) as it is done in the field of reduced sampling techniques or/and by reducing the overall duration of one scan. This is generally achieved by shortening the recovery delay t_{rec} that often dominates the scan duration.

placed in the static magnetic field and being in their thermal equilibrium are manipulated by combining short radio-frequency pulses and delays with the desired information in mind. Their response is then recorded as a function of time over the acquisition period T_{at}. Finally, a recovery period T_{rec} is applied where the excited spins return to their thermal equilibrium before the experiment can be repeated to increase sensitivity or vary delays for experimental reasons.

The overall duration of such a scan of an experiment, T_{scan}, is the sum of the length of the pulse sequence T_{rf}, the acquisition time T_{at} and the recovery delay between two scans T_{rec}. The acquisition time is fixed by the desired spectral resolution, and the length of the pulse sequence by the constraints of the experiment. Also, both durations are rather short, and typically in the order of 100 ms or less. Therefore the most interesting parameter to reduce is the recovery delay that is in the order of seconds ($\approx 5T_1$) often dominating the scan duration.

1.3.2 *Time requirements in multidimensional NMR*

The repetition of scans has particularly relevance in multidimensional (mD) NMR. Introduced by Jeener [65] and Ernst [4] mD NMR correlates spins of different atoms to spread the signal from one-dimension into multiple dimensions. This correlation not only yields an increase in resolution but it also contains valuable information about the protein structure and dynamics [9].

Experimentally, the correlation between individual nuclei is done by transferring magnetization via J-couplings cross scalar bonds, or via dipolar interactions during the mixing period of the pulse sequence followed by labeling of their time evolution depending on the different resonance frequencies. Contrary to 1D NMR it is, generally, not possible to detect all the dimensions directly during the acquisition period. Only one type of nucleus can be detected in the so-called direct dimension. The other dimensions are sampled in an indirect manner by repeating the experiment while step-wise incrementing the evolution delays in the pulse-sequence. They are referred to as indirect dimensions.

A 2D NMR experiment can therefore be described as N repetitions of a scan with different t_1 values to resolve the indirect dimension while the direct dimension t_2 is completely defined over the acquisition period:

$$\left[excitation \Longleftrightarrow t_1 - evolution \Longleftrightarrow mixing \Longleftrightarrow acquisition(t_2)\right]_N$$

This principle is expandable for higher dimensions just by adding mixing and time evolution periods. Figure 3 shows a schematic representation of a 3D experiment correlating three nuclei H-X-Y, and gives an example of sampling on a two-dimensional grid formed by the two time variables of both indirect dimensions t_1 and t_2. The conventional sampling is linear and the difference between two time points is determined by the inverse of the needed spectral width, and the number of sampled points that determines spectral resolution.

It is obvious that data sampling in multidimensional NMR increases dramatically the overall experimental time. A 2D experiment using 50 points to sample the indirect dimension increases the needed scan number by two orders of magnitude, thus several minutes are needed to acquire 2D experiments, several hours for 3D , days for 4D and month for 5D.

In the last ten to fifteen years several experimental concepts have been introduced to shorten experimental time [32]. They can be classified in two major categories. The first category is exploiting different techniques allowing a reduction of the data sampling to minimize the scan number needed to resolve a spectrum [34]. The second category contains fast pulsing methods where the recovery delay between subsequent scans is decreased to a minimum allowing to record more scans in the same time period [143]. Sampling can

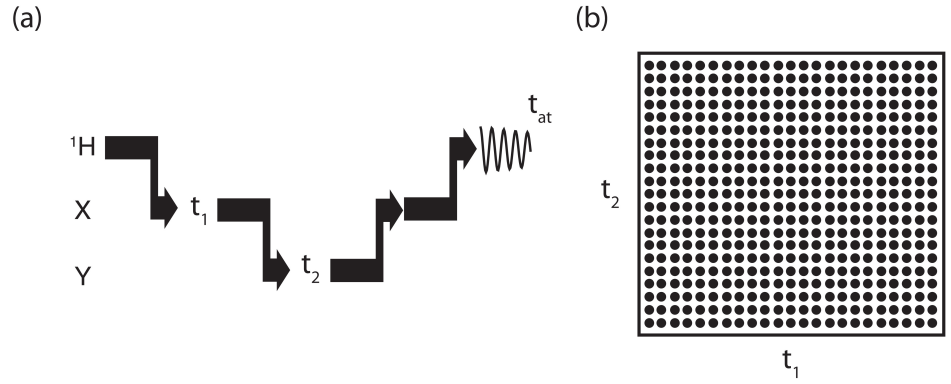


Figure 3 – (a) Schematic representation of three-dimensional NMR experiment with two indirect time dimensions for H-X-Y correlation, t_1, and t_2. (b) Representation of the time domain sampling grip for the two indirect dimensions. Every point in the grid corresponds to one repetition of the pulse sequence for the encoded t_1 and t_2 value.

be reduced by spectral folding [95], projection spectroscopy[73, 85, 166], non-linear data sampling/processing [108, 123], hadamard spectroscopy [34, 169] and spatial encoding techniques [36, 40, 41, 159]. The fast-pulsing techniques use Ernst angle excitation [29, 30], and longitudinal relaxation enhancement [94, 135, 148] by selective excitation of subgroups of protein proton spins to shorten the overall scan time. Both categories are complementary and can be combined together.

The next two sections will briefly cite several examples of ongoing research in the field. Because of the vast number of acquisition schemes and strategies proposed in the last years in this field their aim is to give an overview about what is possible but should not be seen as an exhaustive list. For more information the reader is referred to reviews published by Rovnyak et al. [141] discussing in detail the impact of sampling on the spectral resolution and sensitivity, and Felli et al. [32, 143] for a good overview of different fast NMR techniques.

1.4 FAST PULSING NMR-TECHNIQUES

The concept of fast pulsing NMR is to shorten the recovery delay between subsequent scans where longitudinal magnetization is brought back to its thermal equilibrium by spin relaxation. It allows therefore to collect the same number of scans in less time, for fast spectra acquisition, or to acquire more scans in the same time, for a higher sensitivity and resolution. A main

issue in this approach is that the sensitivity of a following scan depends on the amount of restored magnetization, and decreasing the recovery delay will reduce the amount of available initial magnetization.

Equation 1.1 shows the impact of the used recovery delay T_{rec}, the longitudinal relaxation time constant T_1 describing the time longitudinal magnetization needs to turn back to its thermal equilibrium, and the excitation flip-angle β on the experimental sensitivity.

$$\text{Sensitivity} \propto \frac{(1-\exp(-T_{rec}/T_1))\sin\beta}{1-\exp(-T_{rec}/T_1)\cos\beta}\frac{1}{\sqrt{T_{scan}}} \tag{1.1}$$

It shows that for a flip-angle of $\beta = \pi/2$ the optimal recovery delay is given by $T_{rec} = 1.25\ T_1$. A shorter T_1 relaxation time means therefore a shorter optimal recovery delay. Longitudinal relaxation enhancement (LRE) techniques[135, 144], outlined in the following subsection, are modifying the relaxation behavior of the spin system resulting in shorter T_1 relaxation times allowing fast pulsing. Ernst and his coworkers have also shown that a modification of the flip-angle β can be optimized for the fast pulsing regime to compensate the incomplete magnetization recovery due to the short recycling delay [29, 30]. This will also be discussed in this section.

1.4.1 *Selective excitation of subgroups of protein proton spins to enhance longitudinal relaxation*

Longitudinal relaxation enhancement (LRE) can be accomplished in different ways. One example is the addition of paramagnetic agents enhancing relaxation by their large dipole-dipole moment of unpaired electrons. They can be used to enhance the relaxation either directly on the nuclei or indirectly over the water's ^1H longitudinal relaxation. Another way is an optimized pulse sequence design to achieve faster magnetization recovery. The work presented in this thesis is based on optimized pulse sequence design enhancing longitudinal relaxation by applying selective pulses. Therefore paramagnetic relaxation enhancement will not be further discussed here and the focus will be on a simple description of the principle of selective pulsing.

^1H-^1H dipolar interactions are one dominant mechanism for longitudinal ^1H spin relaxation the so called spin-lattice relaxation. The LRE approach

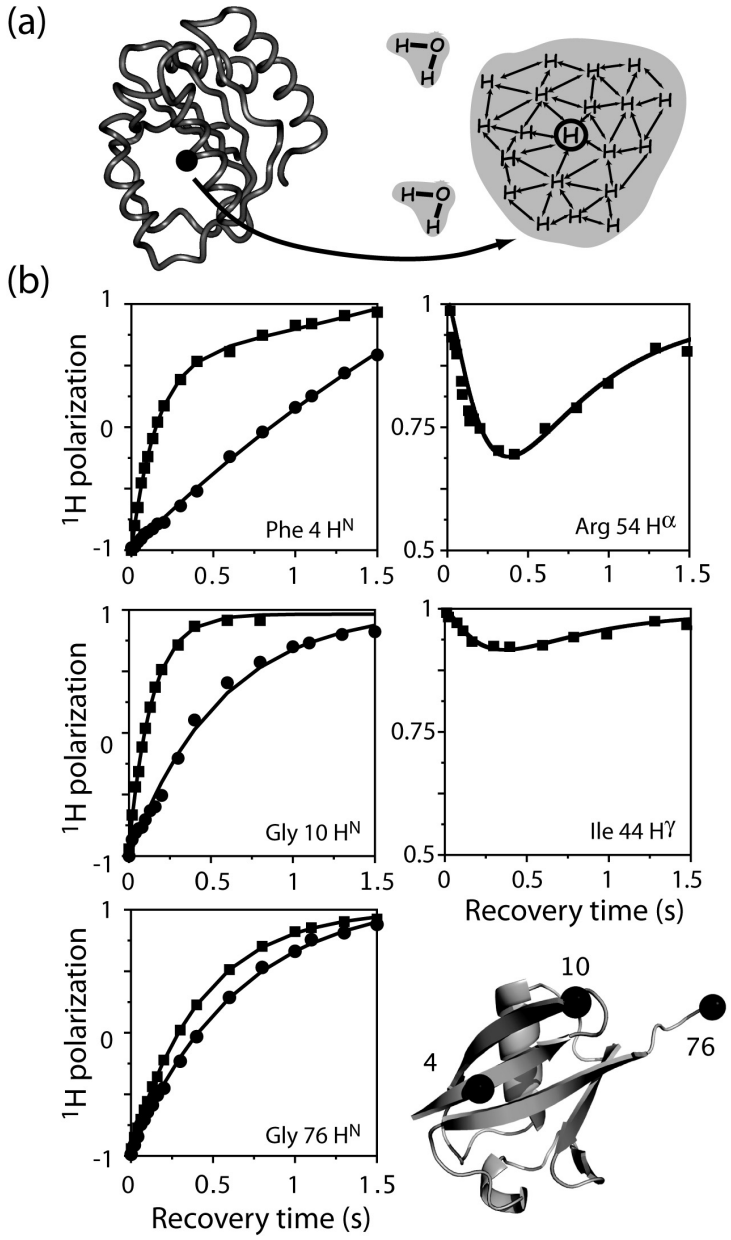


Figure 4 – (a) Schematic representation of the proton network of a protein in solution. Selectively excited protons can relax back to their equilibrium over interactions with environmental protons of the protein and the solvent. (b) Experimental T_1 relaxation curves plotted for 3 residues in ubiquitin. The left side shows the relaxation behavior of the amide proton in the case of non-selective (circles) and selective excitation (squares). The polarization evolution of residual protons kept outside of the excitation band is shown on the right side.

using selective pulses to excite only a subgroup of protons such as amide or aliphatic protons so that the unperturbed protons can take up some energy via dipole-dipole interaction, or hydrogen exchange, results in faster spin-lattice relaxation (figure 4a). The Solomon equations for a two-spin system describing its time evolution can be used to show the influence of selective pulsing on the relaxation of excited and non-excited protons.

$$-\frac{d}{dt}\begin{pmatrix} I_{1z}-I_{1z}^0 \\ I_{2z}-I_{2z}^0 \end{pmatrix} = \begin{pmatrix} \rho & \sigma \\ \sigma & \rho \end{pmatrix}\begin{pmatrix} I_{1z}-I_{1z}^0 \\ I_{2z}-I_{2z}^0 \end{pmatrix} \tag{1.2}$$

I_{1z}^0 and I_{2z}^0 are the thermal equilibrium longitudinal magnetizations, ρ is the auto-relaxation constant and σ is the cross-relaxation constant.

Solving equation 1.2 for the two cases of non-selective ($I_1z = -I_z^0$, $I_2z = -I_z^0$) and selective I_2 inversion ($I_1z = I_z^0$, $I_2z = -I_z^0$) leads to:

non-selective inversion:

$$\begin{aligned} I_{1z}-I_z^0 &= -2I_z^0 \exp^{-(\rho+\sigma)t} \\ I_{2z}-I_z^0 &= -2I_z^0 \exp^{-(\rho+\sigma)t} \end{aligned} \tag{1.3}$$

selective inversion

$$\begin{aligned} I_{1z}-I_z^0 &= -I_z^0(\exp^{-(\rho+\sigma)t}-\exp^{-(\rho-\sigma)t}) \\ I_{2z}-I_z^0 &= -I_z^0(\exp^{-(\rho+\sigma)t}+\exp^{-(\rho-\sigma)t}) \end{aligned} \tag{1.4}$$

The equations for non selective excitation show the classical mono-exponential behavior with a time constant of $(\rho+\sigma)$. In contrast, in the case of selective excitation an additional exponential term is resulting with $(\rho-\sigma)$. While the auto-relaxation term σ is always positive the cross-relaxation term behaves negative for slowly tumbling molecules ($\tau_c > 2ns$). This results in a faster relaxation of the selectively inverted spin while the polarization of the other

spin is slightly decreasing, both spins will then turn back simultaneously to their equilibrium polarization (figure 4b). This is the principle of longitudinal enhanced relaxation used in optimized pulse sequences like the SOFAST [41, 72, 146, 148] or BEST [94] experiments, object of this thesis, where the recovery delays can be shortened from some seconds to mili-seconds.

1.4.2 *Ernst Angle Excitation*

It has been shown that in the case of a single pulse experiment the ideal flip-angle β can be calculated, and that a ninety degree excitation is only optimal in the case of recovering delays of about four times T_1 [29, 30]. The optimal angle in the fast pulsing regime can be written as a function of the recovery time:

$$\cos \beta_{Ernst} = \exp(-T_{rec}/T_1) \tag{1.5}$$

Implementing an Ernst angle optimized excitation pulse in LRE optimized sequences with a dramatically shortened recovery delay leads therefore to significant sensitivity gain. Unfortunately, implementing the Ernst angle into complex sequences using several pulses on the proton channel is rather difficult, but it fits perfectly into sequences using only few proton pulses such as HMQC experiments (SOFAST).

1.5 ACCELERATING SPECTRA ACQUISITION BY REDUCING THE NUMBER OF SCANS

As discussed in the section about multidimensional NMR, the sampling of the indirect dimensions is a very time consuming necessity, thus various techniques to reduce the number of scans needed for spectra acquisition have been proposed. The aim of all of them is to reduce the number of data points collected without significantly lowering the experimental sensitivity and the spectral resolution. Several techniques were therefore proposed to achieve this scan number reduction. They can be classified in two categories. Approaches using an alternative sampling of the indirect dimension grid (sparse non uniform random data sampling, projection NMR, and spectral aliasing) and those replacing the classic time domain sampling by either a frequency domain sampling (Hadamard spectroscopy) or a spatial encoding of evolution frequencies (Spatial encoding). Some

of these techniques can again be combined for optimal performance. This section will give a short overview of the strategies used and their way of sampling the time dimension grid on the example of a 3D experiment, and made recently in this field of current research.

1.5.1 *Random data sampling*

A basic approach consists in sampling the t_1, t_2 pairs in the acquisition grid by selecting them randomly (figure 5). The choice of the number of sampled points depends on the spectral width and needed resolution and can reach up to 10% of the scan number of the experiments linearly sampled counterpart. As the sampling is of a random nature standard fast Fourier transformation, demanding an uniform linear sampling, is no longer possible and other processing techniques are needed to reconstruct the spectra. Therefore all the solutions proposed are dealing with the task to find ways of getting evolution frequencies out of non-linearly sampled data. Most of these algorithms are based on a modeling of the peak position and shape and are numerically adjusting spectral parameters to fit the modeled spectrum to the experimental data.

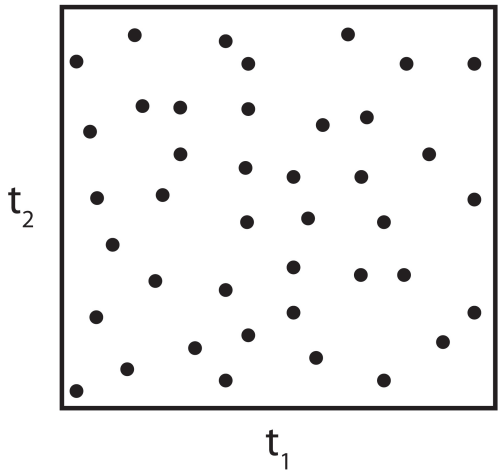


Figure 5 – Example for a random sampling grid for a 3D experiment.

While in the past this kind of approach was very time consuming, enhanced algorithms and modern computers increased significantly the speed of calculation and made random sampling an attractive tool. The main drawback is that it requires either prior knowledge about the spectrum or a trial-and-error adjustment of the fit parameters. Examples of alternative data processing methods are maximum entropy reconstruction [108, 123], the

filter diagonalisation method [107], multi-dimensional decomposition[125], and numerical Fourier integration (NFI) [108]. The latter has the convenient that it does not require modeling of the spectral parameters. All these techniques generally do not need modifications of the regular pulse sequence, and only the non-linear time point sampling has to be implanted.

1.5.2 *Projection NMR or reduced dimensionality NMR*

An alternative to random sampling of the acquired data points is projection NMR where the different time domains are sampled in a non independent way by linking together the step-wise incrementation of different evolution periods. Therefore the approach is also referred to as reduced dimensionality NMR. On the example of a 3D spectrum this means that the two indirect dimensions t_1 and t_2 are replaced by a $t_1 = t$ and $t_2 = \lambda t$ using one time constant t and the projection parameter λ, reducing it to a 2D projection of the spectrum. This results in a radial sampling of the grid as shown in figure 6 where different values are used for the projection parameter λ, defining the projection angle to ($\lambda = \Delta t_1/\Delta t_2 = \arctan \alpha$).

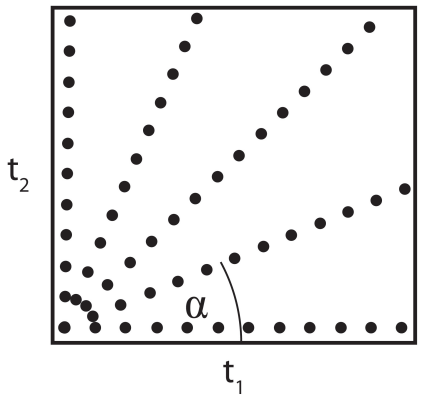


Figure 6 – Example for a sampling grid for a 3D experiment with projection sampling relating t_1 and t_2 by the equations $t_1 = t$ and $t_2 = \lambda t$.

The advantage, contrary to other non-linear sampling approaches, is that it allows, after appropriate recasting, traditional fast Fourier transformation along the linearly sampled trajectories. Transformed spectra show doublets at the sum and the difference of the weighted resonance frequencies of both dimensions ($\omega_1 \pm \lambda\omega_2$), when quadrature detection has been done for one of the frequency-labeled nuclei. In the case of quadrature detection in both

dimensions a single peak is observed centered at either the sum $(\omega_1 + \lambda\omega_2)$ or the difference $(\omega_1 - \lambda\omega_2)$ of the involved resonance frequencies.

Two different approaches are used to extract peak position and intensity out of the reduced dimensionality data. The first approach is based on Fourier transformation along the projection trajectories. Lists of peaks observed in the reduced dimensions are extracted and then individual frequencies and peak intensities are calculated using a linear combination of $\omega_1 \pm \lambda\omega_2$ subsets [73]. The other way is reconstructing the higher dimensional spectrum from a set of projected spectra [84].

1.5.3 *Spectral Aliasing*

If the spectral width of one dimension of a spectrum is reduced below the chemical shift range of the observed nuclei peaks outside the spectral borders are folded back into the spectrum in a predictable way. This is called spectral aliasing and can be exploited to decrease the used spectral width equivalent to an increase of the time increment between two succedent data points (figure 7). This decreases the number of scans collected and therefore the experimental time.

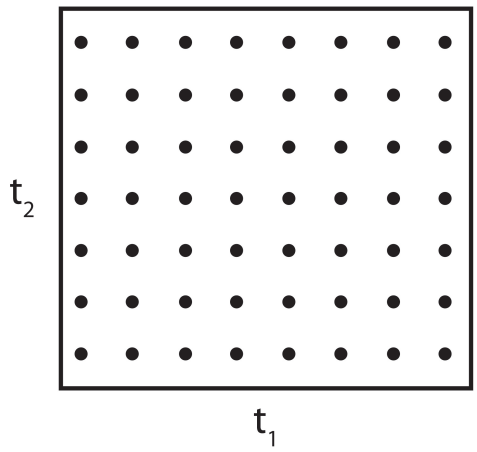


Figure 7 – Example for a sampling grid for a 3D experiment using spectral aliasing sampling.

The optimal spectral width can easily be calculated if one plane of the spectrum is known. A good example of such an application is the set of 3D experiments used for backbone assignment [95]. Based on a ^{1}H-^{15}N HSQC/HMQC spectra it is straightforward to calculate an optimized

spectral width for the ^{15}N-dimension allowing often a dramatic decrease of the experimental time. Because of the simplicity of this approach it can be easily implemented in routine applications after prior chemical shift information is acquired.

1.5.4 *Hadamard spectroscopy*

Hadamard spectroscopy is related to the concept of selective labeling of specific sets of nuclei (more precisely types of amino acids) to simplify the obtained spectra. Imagine a protein containing only a few alanine residues but having a primary structure of about one hundred amino acids. If only the alanine residues are protonated and all the other amino acids deuterated, this allows in most cases to resolve the alanine nuclei in even a 1D spectrum. While specific labeling is getting more and more into the focus of spectroscopists it is still cost intensive and needs several sample preparations with different labeling schemes making it time consuming. Instead of labeling a protein with biochemical means Hadamard spectroscopy uses selective pulses to excite individual resonance frequencies, thus it is also called a discrete frequency-domain sampling (figure 8) [34, 169].

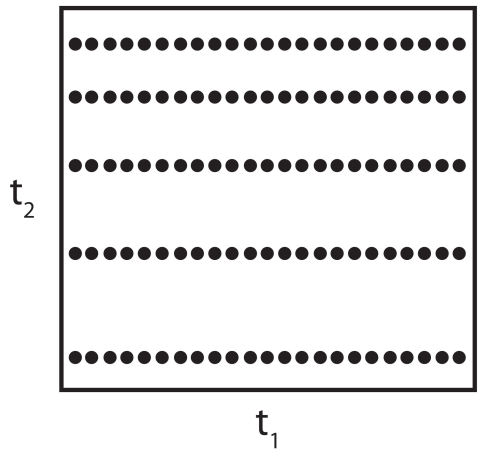


Figure 8 – Example for a sampling grid for a 3D experiment with hadamard data sampling .

If we look at a practical example of n different resonance frequencies in a multidimensional experiment, replacing the usual non selective rf pulse by an array of selective rf pulses, based on prior knowledge of the chemical shift distribution, allows to manipulate the spins resonating at one of the n frequencies individually. In this way sign changes can be created in the

different "frequency slices" of the spectrum. This can be used to replace the traditional chemical shift evolution period t_1 by an Hadamard encoding scheme (m x m Hadamard matrix, where m is equal or higher to the number of resonances n) giving the possibility to separate the different frequencies in the final spectrum. The individual peaks can then be retrieved by summing the scans according to the Hadamard matrix. If the number of frequencies n is lower then the number of sampled points in the classic t_1 approach frequency encoding has a speed advantage, thus making it particularly interesting in the cases of (i) only few peaks in the encoded dimension, and (ii) if the focus is only on some peaks, at example in the active site of a large protein. It is worth mentioning that the resolution of the Hadamard encoded dimension does not depend on the number of scans needed, but on the excitation profile of the selective pulse array used.

An application of Hadamard encoding is the separation of correlation peaks into different 2D spectra depending on the chemical shift of the neighboring ^{13}C atom, being sensitive to the amino acid type [169]. This is especially useful in the context of protein backbone assignment where information about the amino acid type is precious.

1.5.5 *Spatial encoding*

The spatial encoding approach is the most radical approach of the techniques presented in this section. Derived from techniques used in the field of NMR imaging (MRI) to introduce spatial information in the spectrum by the application of pulse field gradients it allows theoretically spectrum acquisition in only one scan [36]. In practice, this technique still suffers some serious drawbacks, where sensitivity and water suppression are the main issues related to the experimental setup needed for spatial encoding. The ultraSOFAST experiment is based on this approach [41] and tackling these experimental problems is one object of the work described in this thesis [39], thus chapter 3 will give a more detailed introduction to spatial encoding.

Part II

RECENT DEVELOPMENTS IN FAST LIQUID STATE NMR PULSE SEQUENCE DESIGN

2

SENSITIVITY ENHANCED IPAP-SOFAST-HMQC FOR FAST PULSING 2D NMR WITH REDUCED RADIO-FREQUENCY LOAD

2.1 INTRODUCTION

Multidimensional nuclear magnetic resonance (NMR) spectroscopy is a widely used tool for high-resolution studies of biomolecular structure and dynamics in solution. While fast time-scale dynamics (ps to ms) are accessible via the measurement of spin relaxation rates [64, 122] or residual dipolar coupling constants [11], kinetic molecular processes taking place on longer time scales can be followed by monitoring NMR spectral changes in real time during a molecular reaction such as protein (un)folding [170, 183]. In the past, atom-resolved real-time NMR studies of kinetic processes in proteins on the seconds, and sub-seconds timescales, have been almost impossible because of the time requirements associated with the recording of multidimensional correlation spectra on macromolecules. In standard 2D NMR only the so-called t_2 time axis is acquired directly, like in 1D NMR, while the additional time axis is monitored as a discrete incrementation of a time variable (t_1) throughout a series of repetitions (scans) of the basic experiment. Hence, a major disadvantage of 2D NMR is the long acquisition time that, regardless of sensitivity considerations, may be needed for obtaining these spectra.

2.2.1 *Standard SOFAST HMQC*

Fast-pulsing NMR experiments provide a means for speeding up multidimensional NMR data acquisition. An prominent example of fast-pulsing NMR is the SOFAST-HMQC experiment introduced recently [144, 148] for recording two-dimensional heteronuclear ^{1}H-^{15}N correlation spectra of proteins in only a few seconds. Short scan times of less than 100 ms are achieved by reducing the recycle delay between scans to almost zero. High sensitivity in this fast-pulsing regime is retained by Ernst-angle excitation [30] and longitudinal relaxation enhancement [135]. All ^{1}H pulses in SOFAST-HMQC are applied band-selectively, leaving the ^{1}H spins outside the excitation band unperturbed. For slowly tumbling molecules ($\omega_H\tau_c \gg 1$) such as proteins, this situation yields efficient energy dissipation from the excited ^{1}H spins to the unperturbed ^{1}H spins via dipolar interactions (spin diffusion). In addition, if the excited spins are amide ^{1}H, amide - water hydrogen exchange also accelerates the return to thermal equilibrium. As a result of these effects, longitudinal ^{1}H relaxation times are reduced from more than a second to typically a few hundred milliseconds [20, 135, 144]. Furthermore, a polychromatic shaped ^{1}H excitation pulse [82] provides the variable flip-angle capabilities required for Ernst-angle excitation that further enhances the sensitivity of SOFAST-HMQC for very short recycle delays. Very good performance of the WATERGATE type element G_1-180° ^{1}H -G_1 is yielded when the water resonance is outside the excited bandwidth. This leads to good water suppression even in one scan permitting very fast spectra acquirement needed to follow kinetics in the order of seconds. An example of average sensitivity as a function of scan time, obtained for ^{1}H-^{15}N SOFAST-HMQC of the small globular protein ubiquitin, is shown in figure 10c and compared to the results obtained using a conventional HSQC-based pulse scheme.

Figure 9a shows the standard SOFAST HMQC experiment. For the sake of simplicity we look at the example of ^{1}H-^{15}N HMQC, but it can also be applied to ^{1}H-^{13}C correlation. It is based on the standard HMQC experiment

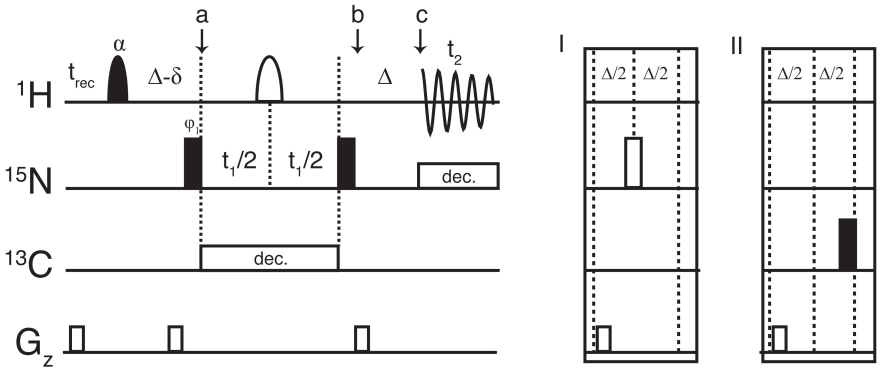


Figure 9 – Standard SOFAST HMQC sequence. (I and II) IPAP filters can be inserted in place of the second Δ delay. Filled and open pulses represent 90° and 180° rf pulses respectively, the first ^{1}H pulse is applied with the flip angle α. The transfer delay is set to $1/(2J_{NH})$ and t_{rec} is the recovery delay between two scans. The δ delay compensates spin evolution during the first selective pulse and has to be adjusted for pure-phase spectra in the direct dimension.

where magnetization is created by an initial 90° rf pulse and then evolving over the coupling period Δ into anti-phase magnetization.

$$H_z \xrightarrow{90°H} H_x \xrightarrow{\Delta} 2H_yN_z \tag{2.1}$$

The first 90° pulse on the nitrogen channel is then creating multiple quantum coherence.

$$2H_yN_z \xrightarrow{90°N} -2H_yN_y \tag{2.2}$$

The nitrogen chemical shift is then evolving over the t_1 period. ^{1}H chemical shift evolution and J-coupling evolution is refocused by the 180° ^{1}H pulse in the middle of the sequence.

$$-2H_yN_y \xrightarrow{t_1} -2H_yN_y\cos(\omega_N t_1) + 2H_yN_x\sin(\omega_N t_1) \tag{2.3}$$

After the evolution period detectable single quantum magnetization is created by the last 90° nitrogen pulse, the multiple-quantum coherence remains undetectable.

$$-2H_yN_y\cos(\omega_N t_1) \xrightarrow{90°N} -2H_yN_z\cos(\omega_N t_1) \tag{2.4}$$

Over the last coupling evolution period the anti-phase component is refocused into transverse proton H_x magnetization prior to detection.

$$-2H_yN_z\cos(\omega_N t_1) \xrightarrow{\Delta} H_x\cos(\omega_N t_1) \tag{2.5}$$

The major advantage of heteronuclear multiple quantum spectroscopy compared to heteronuclear single quantum spectroscopy is that it requires only few rf pulses especially on the proton channel. This has two major advantages for fast pulsing NMR. (i) Limitation of signal loss due to imperfect pulses, and (ii) straightforward implementation of the Ernst angle in the sequence which is rather difficult for more complex sequences.

2.2.2 *IPAP SOFAST HMQC*

In order to achieve the highest possible sensitivity, protein NMR experiments are preferably performed on high-field NMR spectrometers equipped with a cryogenically cooled probe. Such equipment also provides the intrinsic sensitivity required for recording heteronuclear correlation spectra of proteins at sub-millimolar concentrations in a few seconds acquisition time [147]. Generally, duty cycles of a few percent are recommended when running NMR experiments on cryogenically cooled probes to maintain the probe temperature stable at about 25 K. Although it has been shown recently that fast pulsing SOFAST-HMQC experiments with duty cycles on the order of 80% are not incompatible with cryogenic probes [101], the availability of a pulse sequence yielding similar minimal experimental times and comparable sensitivity to SOFAST-HMQC but with reduced radiofrequency (rf) load may still be of interest to reduce rf heating problems, and the risk of long-term damage of expensive cryogenic probes. The main source of rf-heating in SOFAST-HMQC is the composite 15N decoupling applied during 1H detection. Figure 10b in the next section illustrates the kind of effects that may be

observed when performing an ^{15}N-decoupled SOFAST-HMQC experiment with fast repetition rates (Tscan $<$ 100 ms) on a high-field cryogenic probe: First, a slight detuning of the ^{15}N resonance circuit that increases the ^{15}N pulse width by up to 20%. Second, a drop in the detected NMR signal that can be ascribed to an increased B_0 magnetic-field inhomogeneity. A stable situation is reached after about 1 minute of pulsing. Interestingly, these effects are significantly less pronounced when ^{15}N-decoupling is avoided. In the absence of composite decoupling, steady-state conditions are reached after only a few seconds of pulsing. An elegant way to avoid heteronuclear decoupling during ^{1}H detection is the use of an IPAP filter [2, 126] where the 2 doublet components of a scalar (J)-coupled ^{1}H-X spin system are separated in different subspectra. A single correlation spectrum is obtained by adding the 2 subspectra after appropriate shifting of the 2 spectra along the ^{1}H dimension by an amount $\pm$ J/2. This approach has first been proposed and applied to ^{13}C-detected NMR experiments in solids [25] and liquids [7] to remove homonuclear spin coupling-induced line splitting. In the context of the SOFAST experiment this can be achieved by replacing the second coupling evolution period by an IPAP filter (see figure 9).

In Experiment I the ^{1}H-^{15}N coupling is refocused by the 180° ^{15}N pulse. Thus, anti-phase magnetization $2H_yN_z$ is detected.

$$\begin{aligned} & -2H_yN_z\cos(\omega_N t_1) + 2H_yN_x\sin(\omega_N t_1) \\ \xrightarrow{180°N} \; & 2H_yN_z\cos(\omega_N t_1) - 2H_yN_x\sin(\omega_N t_1) \end{aligned} \tag{2.6}$$

In Experiment II, J-coupling is evolving and leading to detected in-phase magnetization H_x while remaining anti-phase magnetization is converted into undetectable multiple quantum coherence by the 90° ^{15}N pulse prior to detection.

$$-2H_yN_z\cos(\omega_N t_1) \xrightarrow{\Delta} H_x\cos(\omega_N t_1) \tag{2.7}$$

This leads to spectra with either in-phase or anti-phase NMR signals in the ^{1}H dimension.

$$R_{IP} : \cos(\omega_N t_1)\cos(\pi J t_2)\exp(i\omega_H t_2) \quad (2.8)$$
$$I_{AP} : \sin(\omega_N t_1)\sin(\pi J t_2)\exp(i(\omega_H t_2 + \pi/2)) \quad (2.9)$$

The two other components needed for quadrature detection are created by phase-cycling of ϕ_1:

$$I_{IP} : \sin(\omega_N t_1)\cos(\pi J t_2)\exp(i\omega_H t_2) \quad (2.10)$$
$$R_{AP} : \cos(\omega_N t_1)\sin(\pi J t_2)\exp(i(\omega_H t_2 + \pi/2)) \quad (2.11)$$

Adding and subtracting the two recorded spectra (R_{IP}+iI_{IP} and R_{AP}+iR_{AP}) leads then to two spectra containing only one of the two ^{1}H doublets. Every spectrum has a 2 times lower S/N noise ratio then a standard SOFAST experiment recorded in the same experimental time. A standard HMQC spectrum can then be obtained by adding the two subspectra after shifting of the two spectra by $\pm$ (J_{HN})/2. This results in an overall sensitivity loss of $\sqrt{2}$ compared to the standard experiment. Another drawback is that the minimum experimental time is two times longer because of the need to detect the in-phase and anti-phase component separately. The next section will show that the IPAP version of the experiment can be significantly optimized yielding the same sensitivity then the standard experiment.

2.3 SENSITIVITY ENHANCED (SE)-IPAP-SOFAST-HMQC

The originally proposed IPAP version of SOFAST-HMQC, however, suffers from a $\sqrt{2}$ sensitivity loss with respect to the ^{15}N-decoupled experiment [148]. Here we introduce a sensitivity-enhanced (SE) version of IPAP-SOFAST-HMQC that provides comparable sensitivity to standard SOFAST-HMQC, and therefore presents an attractive alternative in situations where the NMR experimentalist has to worry about radiofrequency load and duty cycle issues. It also allows the use of much longer t_2 acquisition times in situations where high resolution in the ^{1}H dimension is important. In addition, SE-IPAP-SOFAST-HMQC provides a tool for fast and sensitive measurement of (scalar and dipolar) ^{1}H-^{15}N coupling constants, as demonstrated previously [148].

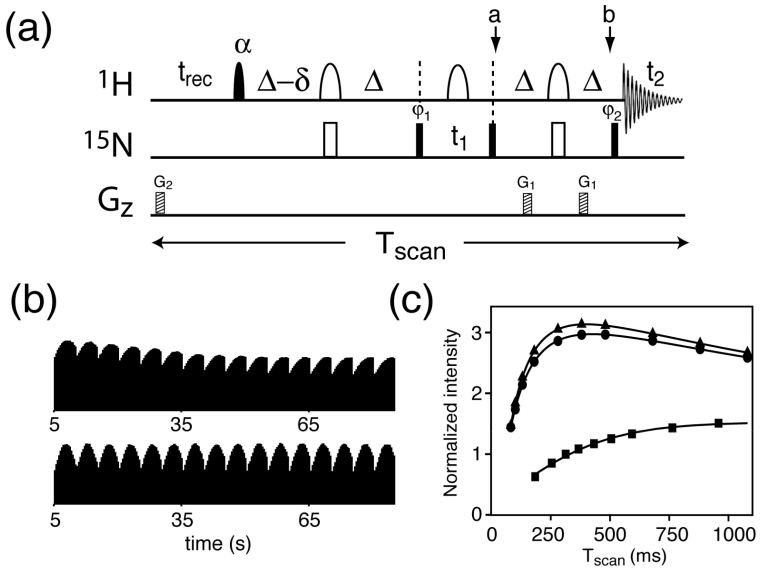


Figure 10 – (a) Pulse sequence for the sensitivity-enhanced (SE) IPAP-SOFAST-HMQC experiment. Filled and open pulse symbols indicate 90° and 180° rf pulses, except for the H^N excitation pulse applied with flip angle α (with α chosen in the range $90° < \alpha < 150°$). The variable-flip-angle H^N pulse has a polychromatic PC9 shape [82], and band-selective H^N refocusing is realized using REBURP [42]. The H^N pulses are typically centered at 8.2 ppm and cover a band width of 4.0 ppm, corresponding to pulse lengths of 2.25 ms (PC9) and 1.52 ms (REBURP) at 800 MHz. The transfer delay Δ is set to $1/(2J_{HN}) \approx 5.4$ ms, and the delay δ accounts for chemical shift and spin coupling evolution during about half of the PC9 pulse. 4 repetitions of the experiment are recorded per t_1 increment with the following phase settings: (1) $\varphi_1 = x$, $\varphi_2 = y$, (2) $\varphi_1 = x$, $\varphi_2 = -y$, (3) $\varphi_1 = y$, $\varphi_2 = y$, (4) $\varphi_1 = y$, $\varphi_2 = -y$. The recorded time domain data sets are then combined as explained in the main text before Fourier transformation. The pulse sequence code for Varian spectrometers and processing protocols for NMR Pipe software are available from the authors upon request. (b) ^{15}N pulse width calibration curves recorded as a function of time for SOFAST-HMQC (upper panel) and SE-IPAP-SOFAST-HMQC (lower panel). The acquisition time was set to t_2^{max} = 50 ms, and low-power WURST-40 [83]] decoupling ($\gamma B_1^{rms}/2\pi = 580$ Hz) was used for the standard version. The integrated spectral intensity is plotted as a function of time. The ^{15}N pulse width was first calibrated setting the recycle delay to t_{rec} = 1s. Then a series of 15 1D spectra was recorded with the ^{15}N pulse width incremented from 80% to 120% of the initially determined optimal value, and setting the recycle delay to t_{rec} = 1 ms. This calibration series, taking about 5s acquisition time each, was repeated for about 2 min (pulsing time) to monitor the effects of rf heating. (c) Average sensitivity of SOFAST-HMQC (triangles), SE-IPAP-SOFAST-HMQC (circles), and SE-HSQC (squares) as a function of scan time. The variable flip angle was set to α =120° for the two SOFAST HMQC experiments. The sensitivity curves were obtained from integrating 2D spectral intensities measured for different recycle delays, normalized for equal experimental times.

The pulse sequence of SE-IPAP-SOFAST-HMQC is shown in figure 10a. To simplify the following discussion we will focus on ^{1}H-^{15}N correlation experiments, although the experiment is equally applicable to other heteronuclear ^{1}H-X two-spin systems, e.g. aromatic ^{13}C-^{1}H and aliphatic $^{13}C^{\alpha}$-$^{1}H^{\alpha}$ groups in proteins, or ^{13}C-^{1}H groups in the sugar and base moieties of nucleic acids. In the standard SOFAST-HMQC sequence only one of the two ^{15}N quadrature components present after the t_1 evolution period is transferred into detectable ^{1}H coherence. This coherence transfer pathway, described by equation 2.12, is preserved in the SE-IPAP-SOFAST-HMQC equence of figure 10a, while a second coherence transfer pathway, given by equation 2.13, also contributes to the detected NMR signal. Chemical shift evolution of the multiple-quantum coherence $2H_yN_x$ present at time point 'a' is refocused by a pair of ^{1}H and ^{15}N 180° pulses. Then an antiphase coherence 2HyNz is created at time point 'b' by an additional ^{15}N 90° pulse.

Pathway (I):

$$\begin{aligned} &\xrightarrow{t_1} 2H_yN_y\cos(\omega_N t_1) \xrightarrow{90°N} 2H_yN_z\cos(\omega_N t_1) \\ &\xrightarrow{2\Delta} H_x\cos(\omega_N t_1) \xrightarrow{90°N} H_x\cos(\omega_N t_1) \end{aligned} \tag{2.12}$$

Pathway (II):

$$\begin{aligned} &\xrightarrow{t_1} 2H_yN_x\sin(\omega_N t_1) \xrightarrow{90°N} 2H_yN_z\sin(\omega_N t_1) \\ &\xrightarrow{2\Delta} 2H_yN_x\sin(\omega_N t_1) \xrightarrow{\pm 90°N} \pm 2H_yN_z\sin(\omega_N t_1) \end{aligned} \tag{2.13}$$

In the absence of ^{15}N decoupling the 2 coherence transfer pathways are detected as inphase (I) and antiphase (II) J_{NH}-doublets during ^{1}H detection. Contrary to the usual planar-mixing-based sensitivity-enhanced quadrature detection scheme [127], the detection of the 2 orthogonal pathways is achieved without additional transfer delays. The 2 pathways can be separated experimentally by adding (I) and subtracting (II) a second data set

recorded with a 180° phase shift of the last 90° ^{15}N pulse (φ_2) yielding NMR signals with the following time modulation functions:

$$R_{IP} : \cos(\omega_N t_1)\cos(\pi J t_2)\exp(i\omega_H t_2) \quad (2.14)$$
$$I_{AP} : \sin(\omega_N t_1)\sin(\pi J t_2)\exp(i(\omega_H t_2 + \pi/2)) \quad (2.15)$$

Additional 90° phase incrementation of φ_1 as usual for quadrature detection in the ^{15}N dimension yields the 2 additional components:

$$I_{IP} : \sin(\omega_N t_1)\cos(\pi J t_2)\exp(i\omega_H t_2) \quad (2.16)$$
$$R_{AP} : \cos(\omega_N t_1)\sin(\pi J t_2)\exp(i(\omega_H t_2 + \pi/2)) \quad (2.17)$$

After complex two-dimensional Fourier transformation (2D FT) of $R_{IP} + iI_{IP}$ and $R_{AP} + iI_{AP}$, and appropriate phasing, 2 spectra are obtained with inphase (IP) and antiphase (AP) J_{NH}-splitting along the ^{1}H dimension, as illustrated in figure 11a. Note that this sensitivity enhancement scheme is similar to sensitivity-enhanced HSQC-type pulse sequences previously proposed for the measurement of heteronculear coupling constants [21, 91], and for multiplet component selection in C^α-C' correlation experiments [90]. The signal-to-noise (S/N) ratio in each of these IP and AP spectra is a factor of 2 lower than in a conventional ^{15}N-decoupled SOFAST-HMQC spectrum recorded under identical conditions because of the observed line splittings in the ^{1}H dimension. Addition and subtraction of these IP and AP spectra separates the upfield and downfield components of the peak doublet into subspectra (figure 11a), and increases the S/N ratio by a factor $\sqrt{2}$. To further increase the signal to noise ratio by a factor $\sqrt{2}$, the two single-transition state spectra are shifted in frequency by $\Delta\nu = J/2$ with respect to each other, and added (figure 11c and b). This is possible because the scalar one-bond J-couplings differ only little from one nuclear site to the other. Overall, if spin-relaxation and pulse-imperfection-induced signal loss can be neglected, the SE-IPAP and standard SOFAST-HMQC experiments are expected to yield the same overall sensitivity for backbone amide groups. For the NH_2 groups of Asn and Gln side chains, the multiple-quantum coherence $2H_yN_x$ present between time points 'a' and 'b' in the sequence of fig. 10a is transformed by the scalar coupling with the additional ^{1}H spin into $4H_yN_yH_z$ that is not converted into detectable ^{1}H coherence by the final ^{15}N 90° pulse. Therefore only pathway (I) is detected for NH_2 groups, and each of the two NH_2 correlation peaks appears as a 1:2:1 triplet motif

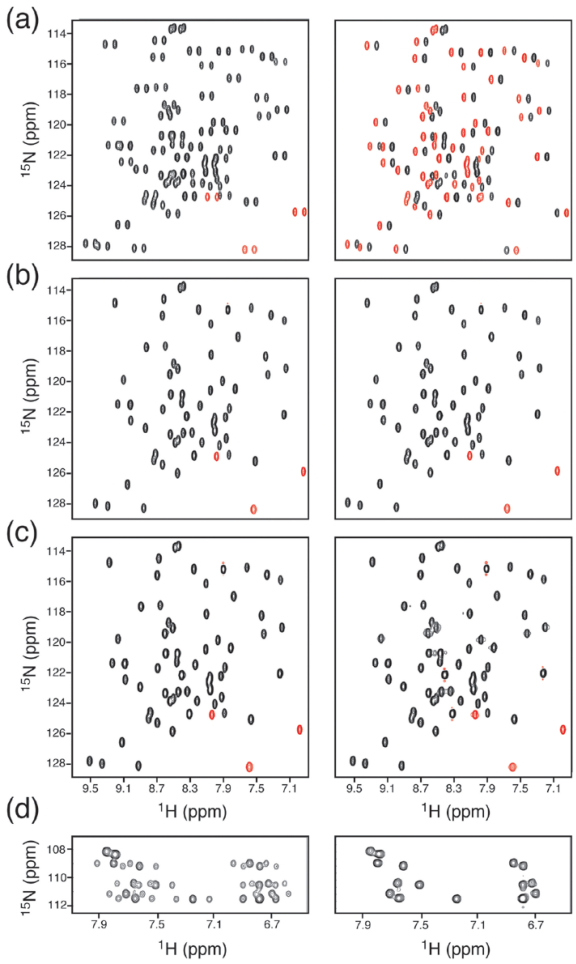


Figure 11 – ^{1}H-^{15}N SE-IPAP-SOFAST-HMQC spectra (central region) of ubiquitin (1.9 mM, pH 6.3) recorded at 25°C on a 800 MHz Varian spectrometer equipped with a coldprobe. The acquisition times were set to t_1^{max} = 28 ms and t_2^{max}= 50 ms, and the recycle delay to t_{rec} = 1 ms, resulting in an experimental time of 30 seconds. The spectra shown correspond to the result of different processing steps as explained in the text. (a) Inphase (left) and antiphase (right) spectra. (b) Single-transition-state spectra, obtained by addition (left) and subtraction (right) of the spectra shown in (a) where the 2 doublet components are detected in different subspectra. These spectra are useful for the measurement of ^{1}H-^{15}N coupling constants. (cb) Final SE-IPAP-SOFAST-HMQC spectrum (left), obtained by adding the spectra shown in (ba) after shifting them by $\pm$ 46.5 Hz with respect to each other along the ^{1}H dimension. A standard SOFAST-HMQC spectrum, recorded on the same ubiquitin sample in the same overall experimental time using identical acquisition parameters, is shown in the right panel for comparison. Additional 1D ^{1}H traces, extracted at a ^{15}N frequency of 115.6 ppm, are plotted for a better appreciation of the S/N ratios. (dc) NH_2 region of the SE-IPAP-SOFAST-HMQC (left) and SOFAST-HMQC (right) spectra.

along the ^{1}H dimension in the final SE-IPAP-SOFAST-HMQC spectrum (fig. 11c). The single-transition state spectra shown in fig. 11a also allow the measurement of the heteronuclear spin coupling constants from the relative peak position in the ^{1}H dimension. This may be of interest for the fast measurement of ^{1}H-^{15}N residual dipolar couplings (RDCs) during the course of a kinetic reaction such as protein folding that can be monitored by real-time 2D NMR spectroscopy [147]. Such kinetic RDC measurements will provide valuable information about structural changes occurring during the kinetic event.

2.4 DISCUSSION AND CONCLUSION

We have experimentally compared the sensitivity obtained for the new SE-IPAP-SOFAST-HMQC and the conventional SOFAST HMQC pulse schemes. The experiments were performed on a Varian Inova 800 spectrometer equipped with a coldprobe using a 1.9 mM sample of ^{15}N-labeled ubiquitin at 25°C. Despite the increased number of ^{15}N rf pulses required for SE-IPAP-SOFAST-HMQC, a comparable overall sensitivity is observed for very short recycle delays using the same number of scans (see figs. 10c and 11b). For longer recycle delays the standard pulse sequence performs slightly better (fig. 10c) mainly because of the smaller number of rf pulses. The maxima of the sensitivity curves shown in fig. 10c are reached at similar scan times (T_{scan}) indicating that the aliphatic and water ^{1}H polarization are only little affected by the additional ^{1}H pulses in the SE-IPAP-SOFAST-HMQC experiment. The minimal experimental time required for SE-IPAP-SOFAST-HMQC, however, is twice as long as for standard SOFAST-HMQC because a total of 4 (instead of 2) repetitions of the basic pulse sequence need to be performed to obtain the different quadrature components (Eqs. 2.14-2.17). Therefore in the same experimental time as required for SE-IPAP-SOFAST-HMQC, it is possible to record a standard SOFAST-HMQC using a two-times longer scan time (T_{scan}). In the fast pulsing regime ($T_{scan} < 100$ ms) this will yield a sensitivity advantage for the standard SOFAST-HMQC version that should be preferred whenever radiofrequency heating is not a problem.

If one has to worry about radiofrequency heating, the SE-IPAP-SOFAST-HMQC pulse sequence provides a sensitive alternative for recording high-quality ^{1}H-^{15}N "fingerprint" spectra of small proteins at sub-millimolar concentration in about 15 s data acquisition time on a high-field NMR spectrometer equipped with a cryogenic probe. The high sensitivity provided by the experimental setup (high-field magnet, cryogenic probe, and

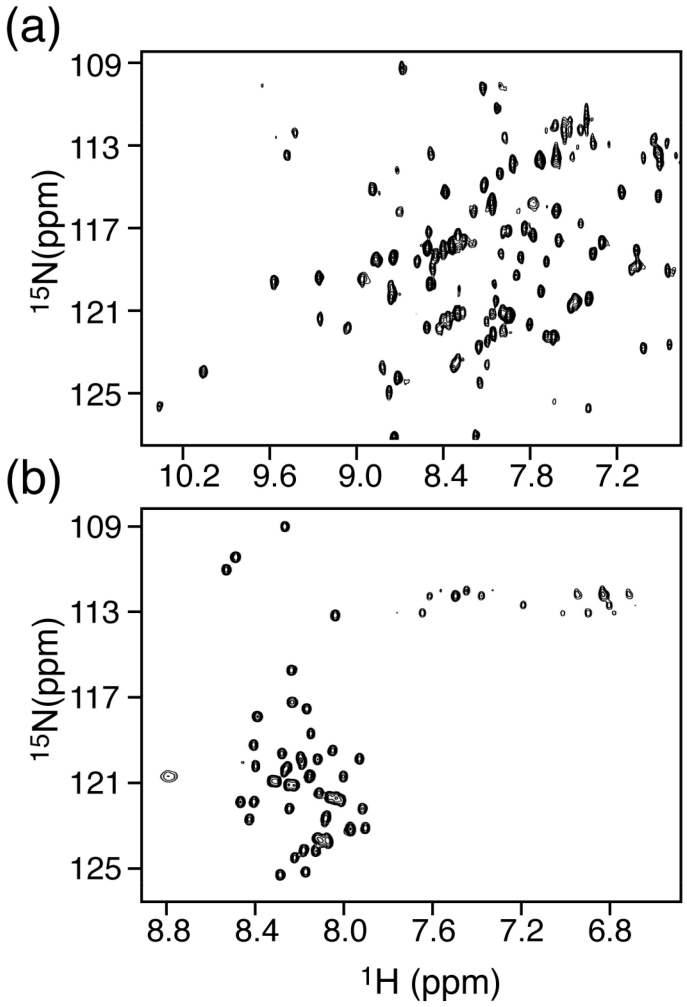


Figure 12 – SE-IPAP-SOFAST-HMQC spectra of 2 protein samples: (a) ^{1}H-^{15}N spectrum of 0.4 mM α-lactalbumin (14 kDa). This spectrum, recorded in only 14 s, demonstrates the high sensitivity provided by the experimental setup. (b) ^{1}H-^{15}N spectrum of a 63-residue highly dynamic α-helical protein. A long acquisition time t_2^{max} = 70ms was chosen for optimal resolution in the detection dimension as most amide ^{1}H resonate within a spectral width of less than 1 ppm.

SE-IPAP-SOFAST-HMQC pulse sequence) allows recording of high-quality ^{1}H-^{15}N "fingerprint" spectra of small proteins at sub-millimolar concentration in about 15 s data acquisition time. This is demonstrated for a 0.4 mM sample of the 14 kDa protein α-lactalbumin in fig. 12a. The absence of ^{15}N-decoupling in SE-IPAP-SOFAST-HMQC allows choosing longer acquisition times in the ^{1}H dimension without increasing the radiofrequency load of the probe. This is especially important when studying helical-rich proteins or partially unfolded systems characterized by reduced chemical shift dispersion in the amide ^{1}H spectrum. An example of a SE-IPAP-SOFAST-HMQC ^{1}H-^{15}N correlation spectrum recorded in 15 s using an acquisition time t_2^{max}= 70 ms on a small highly dynamic α-helical protein is shown in fig. 12b. In conclusion, we have presented a new experiment, SE-IPAP-SOFAST-HMQC that provides comparable sensitivity to standard SOFAST-HMQC while avoiding ^{15}N composite decoupling during ^{1}H detection. Sensitivity enhancement is achieved by a slight modification of the originally proposed IPAP-SOFAST-HMQC pulse sequence that converts both orthogonal transverse ^{15}N components present after t_1 into detectable ^{1}H coherence. SE-IPAP-SOFAST-HMQC provides an attractive alternative to standard SOFAST-HMQC for recording ^{1}H-^{15}N protein spectra in a few seconds, whenever, due to hardware limitations, the use of composite decoupling during detection is not recommended. It also allows the fast measurement of heteronuclear coupling constants. The sensitivity-enhancement "trick" presented here is not limited to SOFAST-HMQC, but it is also applicable to many other IPAP-type NMR experiments. Examples are ^{13}C-detected correlation experiments recently proposed for protonless NMR of proteins and nucleic acids [7, 33] where line splitting due to homonuclear ^{13}C-^{13}C couplings is removed by an IPAP filter.

3

AN IMPROVED ULTRAFAST 2D NMR EXPERIMENT: TOWARDS ATOM RESOLVED REAL TIME STUDIES OF PROTEIN KINETICS AT MULTI-HZ RATES

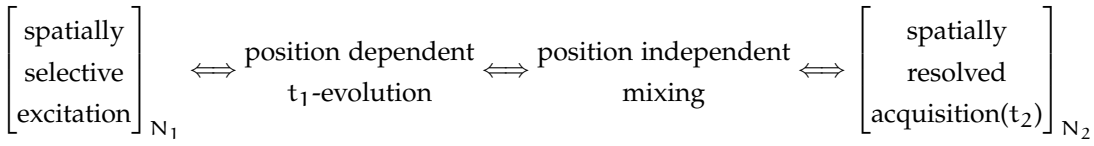
3.1 INTRODUCTION

Several conceptually different approaches have been discussed for acquiring 2D NMR spectra in a significantly reduced amount of time [35, 106, 144]. The most dramatic time reduction is achieved by the so-called "ultrafast" NMR technique that allows the acquisition of 2D (and higher dimensional) NMR spectra within a single scan [36, 159].

3.1.1 *Spatial encoding*

In ultrafast NMR the time incrementation of t_1 is replaced by a spatial encoding step, where nuclear spins in the sample are progressively excited according to their position along a spatial coordinate by the use of a magnetic field gradient acting in combination with a frequency swept radiofrequency (rf) pulse. Figure 13a shows a rather simple way to do this position dependent t_1 labeling. The basic module is consists out of two opposing pulse field gradients and a selective 90° pulse synchronized with the first gradient. The module is repeated N times and the offset of the selective pulse is incremented. The result of this manipulation is a spatial winding of the spin magnetizations with a position-dependent phase $C\Omega_1 z$, where Ω_1 is the resonance frequency of the nuclear spin, C is a spatio-temporal constant depending on the sample length and some user-defined acquisition parameters, relating the Ω_1 evolution in a linear way to the spatial coordinate z. This phase modulation, similar to the $\Omega_1 t_1$ phase behavior in a standard experiment, is conserved during any coherent

mixing sequence required for correlating different nuclear spins in a 2D experiment.



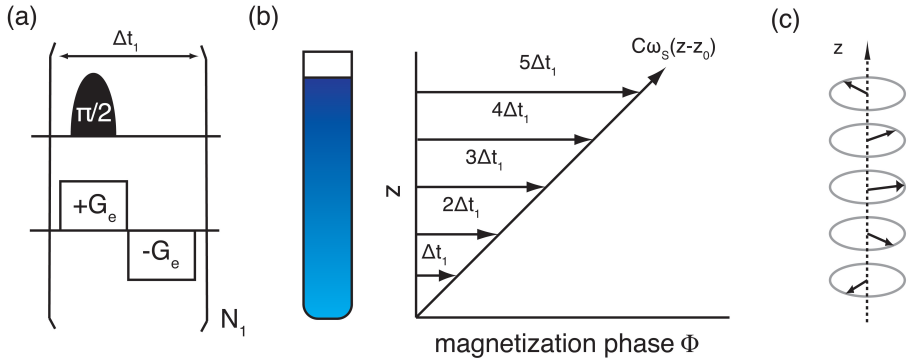


Figure 13 – (a) Example of a pulse sequence to encode an internal frequency Ω_1 along the z axis. Gradient G_e results in a winding along the z axis followed by spatially selective excitation via the shaped 90° pulse and finally rewinding. (b) Schematic representation of spatial frequency dependance created along the sample axis. Here shown on the example of five repetitions of the pulse sequence shown in (a). (c) Schematic representation of the final Ω_1 dependent helical spatial winding preserved after the spatial encoding sequence.

The first applied field gradient is creating a spatial winding of the spins by $\gamma_e G_e z t(z)$ where γ_e is the gyromagnetic moment of the evolving spin during encoding and G_e the magnitude of the applied gradient. The effect of the gradient can be seen like a dispersion of the proper resonance frequency γB_0 proportional to the static magnetic field over a spectral region depending on the gradient strength. This creates a spatial distribution where the resonance frequency of the spin is space-dependant and can be excited separately via the selective 90° pulse. After excitation the effect of the gradient is canceled out by the opposite gradient, resulting in a final evolution phase $\phi(z)$ reflecting only the internal spin frequency and its position. Taking the example of a five time repetition of the excitation module the result

will be 5 subsamples having evolved during t_1 evolution times going from Δt_1 to $5\Delta t_1$ in time steps of Δt_1 like shown in figure 13b. This can be represented in a chemical shift depending spatial-winding (figure 13c). The final accumulated phase $\phi(z)$ after the encoding block will is a pure function of the chemical shift Ω_1 and the spatial position z and can be described as a discrete evolution phase.

$$\{\phi_j(z) = C\Omega_1(z_j - z_{N_1})\}_{j=0,j=N_1-1} \tag{3.1}$$

3.1.2 *spatial decoding*

A spatial decoding is achieved during data acquisition, by the use of a magnetic field gradient. This acquisition gradient G_a with the same spatial dependency then the encoding gradient G_e results in a constructive interference of the spins along the spatial axis of the sample, a spin echo whose position in time along the $k = \gamma \int G_a(t)\,dt$ axis depends on the resonance frequency Ω_1 as well as on the strength of G_a. This reading process can be repeated numerous times by oscillating the sign of the readout gradients, thus allowing to monitor a set of indirect time-domain spectra as a function of the t_2 time evolution. Fourier transformation along the t_2 dimension then results in the desired 2D NMR spectrum. The acquisition module is shown in figure 14a. Assuming the phase has been preserved the signal would be

$$S(k, t_2) \propto \left[\sum_{j=0}^{N_1-1} e^{iC\Omega_1(z_j - z_{N_1})} e^{ikz_j}\right] e^{i\Omega_2 t_2} \tag{3.2}$$

where Ω_2 is the chemical shift acting during acquisition time t_2 and k is the wavenumber associated to the acquisition gradient G_a. It will oscilate between $\pm G_a$ N_2 times during acquisition.

$$k = \int_0^{t_2} \gamma_a G_a(t)\,dt \tag{3.3}$$

The bracketed part of the equation 3.2 is a Fourier transformation giving rise to a signal only when $k = -C\Omega_1$ as the chemical shift induced winding is

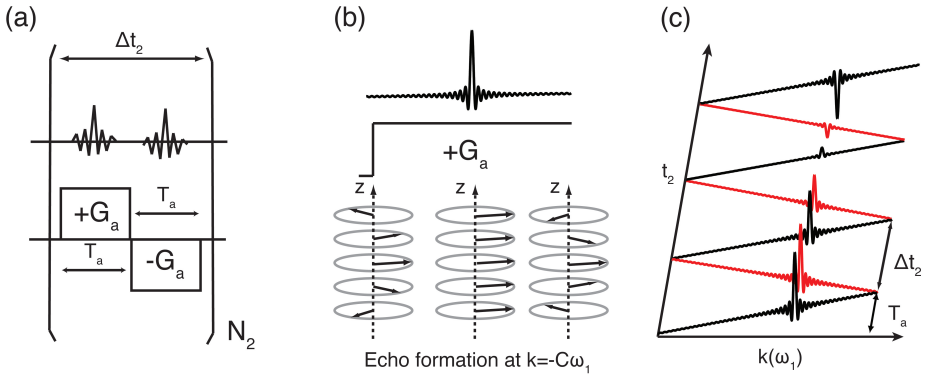


Figure 14 – (a) Gradient echo-module to detect the spatial encoded spin frequencies. The module is repeated N times to create N effective data points along the t_2 axis with a resolution of $\Omega_2 t_2$. (c) Creation of an echo at $k = -C\Omega_1$ in presence of the acquisition gradient G_a (b) Schematic representation of data points acquired while positive (black) and negative (red) gradient acquisition. Echos are time modulated in t_2 by Ω_2

unwinded by the applied acquisition gradient G_a creating overall constructive interference (figure14b). In other words the peak position in the indirect dimension is not given by a FT of a time-evolving coherence, but by the position of the echo in the k space created by the interfering magnetization from different slices.

By this alternation of positive and negative acquisition gradients over a period of t_2 with a period of Δt_2 ($2*T_a$) the signal is acquiring an evolution phase.

$$\phi = kz + t_2\omega_2 \tag{3.4}$$

This gives rise to the in figure 14c represented zig-zaging in the k space where the position of the echo is related to the initial chemical shift Ω_1 and the echo intensity is modulated by t_2. The echos of the positive and negative acquisition gradient data sets can then be separated to correct the "shearing" in the k space and Fourier transformed along the t_2 dimension to result into a classic 2D NMR (Ω_1, Ω_2) spectrum.

3.1.3 *Technical aspects of spatial decoding experiments*

As the setup of a single scan experiment is very different to a standard t_1 time domain encoded experiment some experimental parameters will be briefly discussed. Taking the acquisition gradient $\pm G_a$ with a duration time of T_a and a physical dwell time of Δt, the number of complex points collected along the k-axis will be $N_k = T_a/\Delta t$. When the acquisition cycle is repeated N_2 times the overall acquisition time is $2T_aN_2$ with $2N_kN_2$ complex point being collected. Points according to the same gradient set (positive or negative acquisition gradient) are spaced by T_a this results into a spectral width of:

$$sw_2 = \frac{1}{2T_a} \tag{3.5}$$

The maximum k value is determined by $k_{max} = \gamma_a G_a T_a$ and directly related to the spectral width $sw_1 = k_{max}/C$ with $C = \Delta t_1/\Delta_z = \Delta t_1/\left[\Delta O/\gamma_e G_e\right]$, where ΔO is the increment of the offset of the selective excitation pulses. If we assume that $\Delta t_1 = 2T_e$ we can write

$$sw_1 = \Delta O \frac{\gamma_a G_a T_a}{2\gamma_e G_e T_e} \tag{3.6}$$

Being maybe the only technique allowing the acquisition of a whole nD spectrum in one scan its biggest disadvantage is its reduced sensitivity. An important aspect is the gradient based encoding dispersing the spin frequencies over a broad spectral width. An opening of the filter bandwidth over several ten thousands of hertz becomes necessary.

$$fbw = \gamma_a G_a L/2 \tag{3.7}$$

This goes in hand with an increase in the collected noise that is directly proportional to the filter bandwidth. Thus single-scan spectra acquisition suffers from significant sensitivity loss due to increased noise detection.

3.1.4 *UltraSOFAST*

Recently the two complementary ideas, recording a 2D data set in a single scan by spatial frequency encoding, and reducing the repetition rate with-

out compromising sensitivity by enhanced longitudinal ^{1}H spin relaxation [135] have been combined in the ultraSOFAST-HMQC experiment shown in Fig. 15 [41]. Although this experiment allowed us to demonstrate the feasibility of acquiring a series of protein ^{1}H-^{15}N spectra at Hz rates, an excess of signal loss related to different aspects of the experiment discussed below was observed, that limits its applicability in the context of real-time studies of molecular kinetics. Moreover, for the short repetition times used for this experiment, the rf load and the pulsed field gradients required for spatial encoding and decoding could easily exceed the duty cycle recommended for an optimum performance of the probe, particularly in the case of cryogenically cooled probes. Challenged by these features, we have developed, single-transition-state (STS)-ultraSOFAST-HMQC, an experiment that has the same characteristics in terms of minimal experimental times and repetition rates as ultraSOFAST-HMQC, but that overcomes some of the major limitations of the latter. STS- ultraSOFAST-HMQC yields increased sensitivity under fast-pulsing conditions when compared to ultraSOFAST-HMQC. This is demonstrated by the measurement of fast D/H exchange kinetics from a real-time series of ^{1}H-^{15}N spectra recorded at a rate of ≈ 3 Hz.

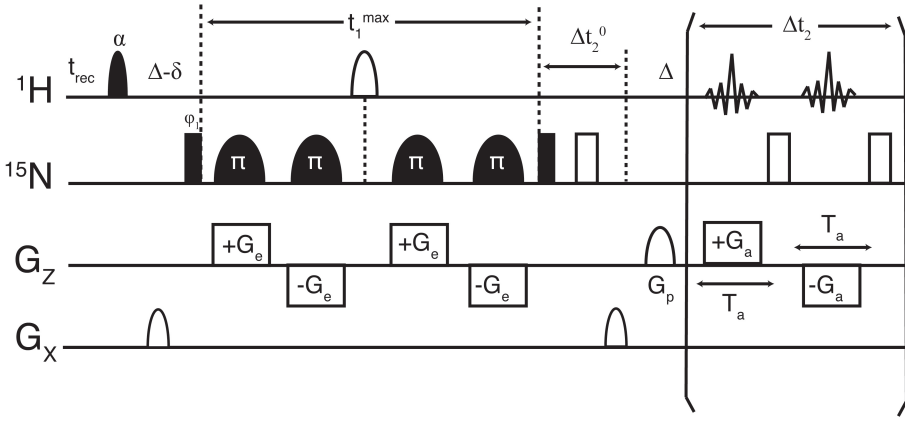


Figure 15 – NMR pulse sequence for 2D ^{1}H-^{15}N UltraSOFAST HMQC. Filled and open pulses indicate 90° and 180° rf pulses, respectively, except for the first ^{1}H pulses that is applied with a variable flip-angle for Ernst angle excitation.

Ultrafast 2D NMR allows already to acquire a spectrum in one scan and one would think that fast repetition rates of the experiment are not of greater use because only one scan is needed. Nevertheless, although it is in most of the cases necessary to acquire at least a minimum number of scans to obtain sufficient sensitivity and solvent suppression. Also, when following

dynamical processes like protein folding or D/H exchange it is necessary to repeat the experiment over a longer time period to monitor spectral changes as a function of time. Therefore the ultraSOFAST approach enables fast repetition and higher sensitivity, crucial for this kind of studies.

The implementation of the Ernst-Angle is done on the selective 90° ^{1}H excitation pulse having a flip-angle α higher than 90°. All ^{1}H pluses applied are band-selective on the subset of protons of interest. In the case of the sequence shown in figure 15 this are the amide protons. All other protons are left in equilibrium allowing fast T_1 relaxation of the amide protons to the thermal equilibrium and therefore short repetition rates. The spatial encoding is implemented by a pair of bipolar gradients and π sweep pulse instead of the previously shown repetition of selective 90° pulses. This allows a constant time kind labeling. Combined with the 180° ^{1}H pulse in the middle of the pulse sequence it permits cleaner refocusing of the coupling and the gradient effects of the ^{1}H evolution while the spatial encoding. Decoupling of the scalar coupling during acquisition is done by 180° ^{15}N hard pulses between each acquisition gradient.

Sensitivity in ultraSOFAST NMR

One important point in ultrafast NMR is the sensitivity. It comes from the nature of spatial encoding itself that sensitivity is a big issue. Naturally a very good intrinsic sensitivity is needed to obtain sufficient signal in one scan and therefore high protein concentrations in the lower millimolar (2-3 mM) range are often needed. But there are also sever sensitivity losses due to translational diffusion and pulse imperfections during the encoding module. Another origin for a lower SNR is the decoupling during acquisition. As previously discussed the receiver bandwidth has to be opened at about $\gamma_a G_a L/2$ due to the gradient encoding. The spectral bandwidth in the carbon dimension is linearly proportional to the product of the gradient strength and its duration $G_a T_a$. The presence of the ^{15}N decoupling pulses is reducing it by the length of the pulse and the delay between the pulse and the gradient. This means that the gradient strength has to be increased to keep the same spectral width. Given that the SNR is proportional to $\sqrt{G_a}$ it is in this case also directly related to its duration T_a. Thus, decoupling is reducing considerably SNR and a non-decoupled version of the sequence would greatly enhance the sensitivity of the experiment.

Spatially encoded ultrafast NMR experiments generally result in decreased sensitivity when compared to their standard time-evolution based counterparts. Neglecting sensitivity drops arising from non-ideal radiofrequency and gradient pulses, most of these losses can be rationalized by the following: NMR data acquisition in the presence of a magnetic field gradient requires opening of the filter bandwidth to account for the gradient-induced spread in resonance frequencies from spins over the whole sample volume. Given that the noise detected by the receiver increases with the receiver's bandwidth, the overall sensitivity of the experiment strongly depends on the strength of the acquisition gradients - dictated in turn by spectral widths considerations (See results). As a consequence, while high magnetic fields provide the high spin polarization and frequency resolution required for (ultra)fast protein NMR methods, the intrinsic per-scan sensitivity of an ultrafast experiment is lower then that of its conventional counterpart.

Different approaches have been proposed in the past to increase this sensitivity aspect of ultrafast NMR, including optimized data acquisition and processing [40]. Here we focus on a different source of sensitivity losses, arising from problems related to ^{15}N decoupling during the ^{1}H detection period. Due to the large effective spectral spreads brought about by oscillatory magnetic field gradients during acquisition, heteronuclear decoupling in ultrafast NMR experiments is generally not performed using continuous decoupling schemes. Rather, it is realized by 180° high-power refocusing pulses applied at rates SW_H or $2SW_H$, with SW_H the spectral width in the ^{1}H dimension. The finite length of these decoupling pulses, as well as a short recovery delay needed prior to each such pulse, have a direct influence on the required acquisition gradient strength as they reduce the time available for applying a gradient pulse. Owing to this reduced duration, the gradient strength has to be increased in order to cover the same effective spectral width along the spatially encoded dimension. The filter bandwidth, in turn, has to be larger resulting in a further increase in the detected noise and thus a reduction in sensitivity. ^{15}N decoupling during ^{1}H detection is also the main source of rf probe heating leading to partial probe detuning, increased B_0-field inhomogeneity, and the risk of long-term damage of expensive cryogenic probes [72]. Due to all these reasons combined, the development of alternative techniques that do not require heteronuclear decoupling during acquisition is of particular interest for ultrafast NMR experiments.

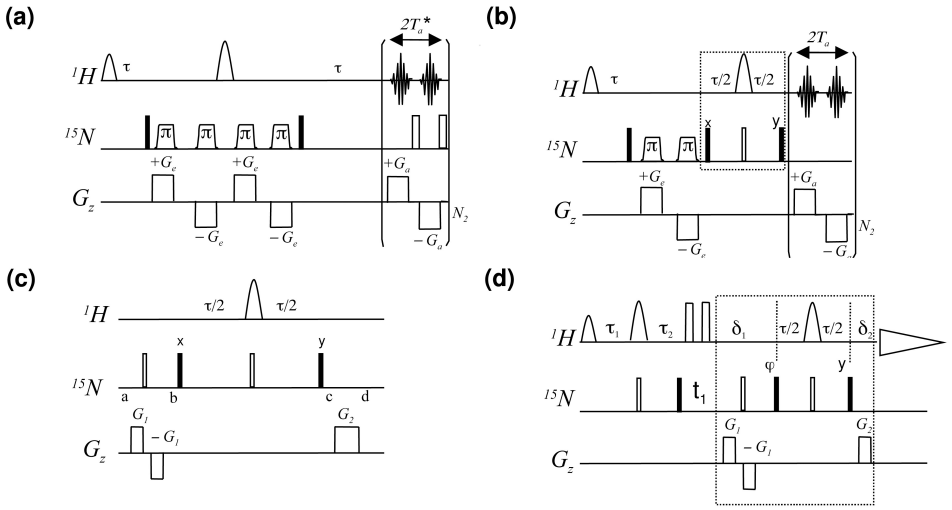


Figure 16 – NMR pulse sequences used in this study for performing fast 2D ^{1}H-^{15}N correlation experiments of proteins: (a) UltraSOFAST-HMQC, (b) STS-ultraSOFAST-HMQC, (c) STS-SOFAST-HMQC, and (d) the STS pulse sequence element. Filled and open pulse symbols indicate 180° and 90° rf pulses, respectively, except for the first ^{1}H pulse that is applied with variable flip angle adjusted for Ernst-angle excitation [144]. Unless indicated, all pulses are applied along the x axis. The frequency-chirped pulses used for spatial encoding in the ultrafast sequences (a) and (b) are achieved by WURST-like amplitude modulation [83] suitable for executing 180° nutations [132]. All selective ^{1}H pulses are typically centered at 8.2 ppm, covering a bandwidth of 4.0 ppm, with the following shapes: PC9 [82] for excitation, and REBURP [42] for refocusing purposes. In all experiments with more than one scan, a simple 2-step phase cycle was applied by inverting the phase of the 90° ^{15}N excitation pulse in concert with the receiver phase. In sequence (c) a pair of ^{1}H 180° broadband inversion pulses [160] is applied during the t_1 evolution period. These pulses refocus ^{1}H chemical shift evolution during t_1 without significantly perturbing the polarization of non-amide ^{1}H. The transfer delays τ_1 and τ_2 are set to $\tau_1 = \tau/2 + \delta_1 - \delta_2$ and $\tau_1 = \tau/2 - \delta_1 + \delta_2$. For Ernst angle excitation, a flip angle $\alpha > 90°$ (typically 120° to 150°) is required for sequences (a) and (b), whereas $\alpha < 90°$ (typically 40° to 60°) is chosen for sequence (c).

An alternative to heteronuclear decoupling involves the use of spin-state-selective NMR techniques, creating a single-transition spin state that evolves during the final detection. This results in a single resonance line shifted either upfield or downfield by $\Delta\nu = J_{NH}/2$ with respect to the standard ^{15}N-decoupled spectrum without the need for implementing an active ^{15}N decoupling. Examples of spin-state selection techniques include the S^{3}E [116], double-S^{3}CT [117, 133], and IPAP-type [2, 126] filters. The S^{3}E and double-S^{3}CT sequences require additional 90° rf pulses on the ^{1}H channel, and they are thus incompatible with the SOFAST-HMQC experiment. IPAP is compatible with SOFAST-HMQC but it doubles the minimal experimental time as 2 spectra need to be recorded yielding inphase, and anti-phase doublet lines, respectively. A singlet spectrum is only obtained after appropriate combination of the two data sets as described recently for IPAP versions of SOFAST-HMQC [72, 144]. Here we present a new single-transition-state (STS) filter sequence that yields spin-state selection in the ^{1}H dimension within a single scan. This STS filter element, depicted inside the square line in Fig. 16b, is compatible with both, spatial encoding and SOFAST-HMQC and can therefore be employed without affecting the minimal achievable 2D frame rate.

To better understand the operation of this new approach to coupling-free 2D NMR, we focus first on the STS filter element highlighted in Fig. 16c. Assuming that an initial multiple-quantum spin state $\sigma_a = 2H_yN_x\cos(k_1z) + 2H_yN_y\sin(k_1z)$ has been created at time point a, and neglecting chemical shift evolution, the gradient pulses G_1 induce a winding of the spin polarization along the z-axis yielding a spin state at point b given by:

$$\sigma_b = 2H_yN_x\cos(k_z) + 2H_yN_y\sin(k_1z) \tag{3.8}$$

with the wave number $k_1 = 2\gamma_N\int_0^{\tau_1} G_1(t)\,dt$, and τ_1 the gradient length. Between time points b and c, the 2 coherences, $2H_yN_x$ and $2H_yN_y$ then evolve along separate pathways as:

$$\begin{aligned} &(I) \quad 2H_yN_y \xrightarrow{90°N_x} 2H_yN_z \xrightarrow{2\Delta} H_x \xrightarrow{90°N_y} H_x \\ &(II) \quad 2H_yN_x \xrightarrow{90°N_x} 2H_yN_x \xrightarrow{2\Delta} 2H_yN_x \xrightarrow{90°N_y} 2H_yN_z \end{aligned} \tag{3.9}$$

After the final gradient pulse G_2 with wave number $k_2 = \gamma_H \int_0^{\tau_2} G_2(t)\,dt$, the spin state is given by:

$$\begin{aligned}\sigma_d &= \{H_x \cos(k_1 z) + 2H_y N_z \sin(k_1 z)\} \exp(ik_2 z) \\ &= \{H_x \exp[i(k_1 + k_2)z] + H_x \exp[i(-k_1 + k_2)z]\}/2 \\ &\quad + \{H_y N_z \exp[i(k_1 + k_2)z] - 2H_y N_z \exp[i(-k_1 + k_2)z]\}/2i \\ &= (H_x + 2H_x N_z) \exp[i(k_1 + k_2)z]/2 + (H_x - 2H_x N_z) \exp[i(-k_1 + k_2)z]/2\end{aligned} \tag{3.10}$$

The $(H_x + 2H_x N_z)$ and $(H_x - 2H_x N_z)$ operators describe single-transition spin states, corresponding to the upfield and downfield ^{1}H doublet components. Experimentally, one of the doublet components is selected by setting the relative gradient strengths to either $k_2 = -k_1$ or $k_2 = k1$. The unwanted doublet component stays dephased along the z-axis, and averages to zero when integrating over the whole sample length.

The STS filter element of Fig. 16c can be inserted into the standard SOFAST-HMQC experiment, as well as into the spatially encoded ultraSOFAST-HMQC. In the STS-SOFAST-HMQC experiment of Fig. 16d, the STS block is inserted after the t_1 evolution period. As a consequence, the accumulated phase at time point b is the combination of the ^{15}N chemical shift evolution and the spatial winding. $\Omega_N t_1 + k_1 z$. Assuming $k_2 = -k_1$, the spin state at time point d is then given by:

$$\begin{aligned}\sigma_d^I(t_1) &= (H_x + 2H_x N_z) \exp(i\Omega_N t_1)/2 \\ &\quad + (H_x - 2H_x N_z) \exp(-i\Omega_N t_1) \exp(-i2k_1 z)/2\end{aligned} \tag{3.11}$$

Only the first part of Eq. 3.11, corresponding to one of the ^{1}H doublet components, yields a detectable signal that is phase modulated as a function of the ^{15}N chemical shift and the incremented t_1 delay. The second part of the density matrix, corresponding to the other ^{1}H component, is dephased along a spatial axis, and does not yield any detectable magnetization. As usual for echo/anti-echo quadrature detection schemes [127], a second experiment is required that yields a negative phase evolution in the t_1 dimension:

$$\begin{aligned}\sigma_d^{II}(t_1) &= (H_x + 2H_x N_z) \exp(-i\Omega_N t_1)/2 \\ &\quad + (H_x - 2H_x N_z) \exp(i\Omega_N t_1) \exp(i2k_1 z)/2\end{aligned} \tag{3.12}$$

This is realized experimentally by inverting the gradient G_2 ($k_2 = k_1$), as well as phase ϕ of the first 90° ^{15}N pulse in the STS element.

In the STS version of the ultraSOFAST-HMQC experiment (Fig. 16b), the G_1 gradient pulses are replaced by a constant-time spatial-encoding scheme. As a result of the combined action of the pulsed field gradients G_e and the frequency-swept ^{15}N pulses, the wave number of the effective phase encoding at time point b is given by $k_1 = C\Omega_1$, with $C = t_{1max}/L$, and L the effective sample length. During the acquisition gradient, data points are sampled as a function of a wave number $k(t) = \gamma_H \int_0^t G_a(t')\,dt'$ yielding a readout spin state given by:

$$\begin{aligned}\sigma(k) &= (H_x + 2H_xN_z)\exp[i(C\Omega_N - k(t))z]/2 \\ &\quad +(H_x - 2H_xN_z)\exp[i(k(t) - C\Omega_N)z]/2\end{aligned} \tag{3.13}$$

For the standard ultraSOFAST-HMQC experiment (Fig. 16a) the same time dependence is obtained for spin evolution during acquisition but only the first coherence-transfer pathway of Eq. 3.9, yielding inphase coherence, contributes to the detected signal. Therefore, from a theoretical point of view, and neglecting the consequences of finite ^{15}N pulse lengths, the standard and STS versions of the ultraSOFAST-HMQC experiment are supposed to provide the same signal intensity-yet freeing the experiment from the need of active ^{15}N decoupling.

3.3 RESULTS AND DISCUSSION

3.3.1 *STS-ultraSOFAST-HMQC*

An experimental comparison between the performance of a standard (Fig. 16a) and STS (Fig. 16b) version of ultraSOFAST-HMQC 2D NMR is shown in Fig. 17. Let us emphasize again that no additional sensitivity loss is expected from the single-transition-state selection in the STS-ultraSOFAST-HMQC experiment. All data sets were recorded on an 800 MHz spectrometer equipped with a cryogenically cooled probe on a 2.3 mM sample of ^{15}N labeled ubiquitin, using identical acquisition parameters except for the acquisition gradients, and for the ^{15}N decoupling procedures. The spectrum in panel (a) was obtained using the standard ultraSOFAST-HMQC experiment (Fig. 16a), where ^{15}N decoupling was realized by high-power 180° pulses

placed within $\pm G_a$ transitions and a 10 µs delay was inserted between the G_a gradient and the ^{15}N pulses. In the spectrum of panel (b) this delay was increased to 20 µs; no longer recovery delays were tested, as this would have increased the acquisition gradient strength to a level too high for the experimental setup. The spectrum in Fig. 17c recorded using the STS pulse of Fig. 16b does not require any ^{15}N decoupling. For a fair comparison of spectral line shapes and signal intensities, no post processing such as linear prediction or signal apodization was employed.

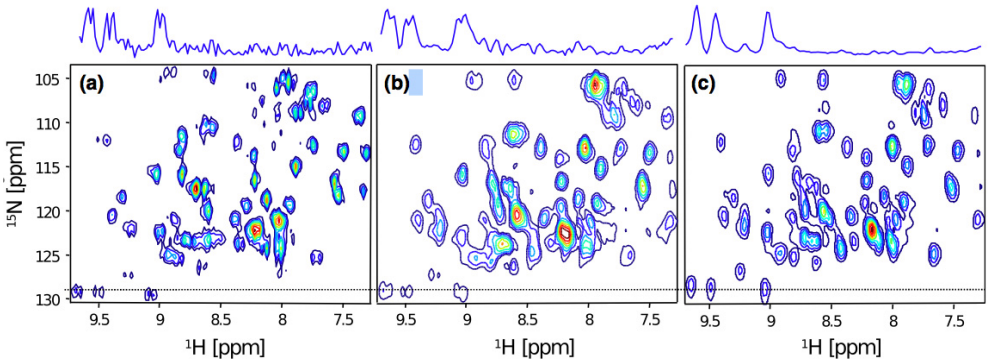


Figure 17 – Sensitivity comparison of standard versus STS versions of ultraSOFAST-HMQC. All spectra were acquired on a 600 MHz spectrometer equipped with a cold-probe using the following experimental parameter settings: $G_e = 20$ G/cm, $N_2 = 64$, $t_1^{max} = 10$ ms, $T_a^* = 250$ µs, ^{15}N 180° pulse length = 64 µs, G_a =29.3, 33.2 G/cm and 20.8 G/cm for panels a,b and c, respectively. The recovery delay between acquisition gradient and ^{15}N 180° pulses was set to (a) $\epsilon = 10$ µs, and (b) $\epsilon = 20$ µs

A significantly improved spectral quality and signal to noise ratio (SNR) are observed when comparing the STS spectrum against the spectra recorded with the standard ultraSOFAST-HMQC version. The spectra in Fig. 17a and b, display a residual coupling of about ≈ 35 and ≈ 20 Hz, respectively, along the ^{1}H dimension. This is most-likely resulting from magnetic field currents that extend into the $+G_a \leftrightarrow -G_a$ switching times, and spoil the efficiency of the decoupling pulses. The use of a longer delay between the PFG and ^{15}N pulses increases the decoupling performance, but further decreases the sensitivity of the experiment (Fig. 16b). This can be explained by the dependence of the SNR on the strength of the acquisition gradient G_a and its dependence on the duration T_a. In the ultraSOFAST-HMQC experiment the effective spectral width in the F_1 (^{15}N) and F_2 (^{1}H) dimensions are given by $SW_N(F_1) = \frac{\gamma_H G_a T_a L}{2t_1^{max}}$ and $SW_H(F_2) = \frac{1}{2T_a^*}$, where $T_a^* = T_a + \delta N + \epsilon$ is half the period of a full acquisition loop including an acquisition gradient, the duration of the including an acquisition gradient, the duration of the ^{15}N π

pulse (δN), and the gradient recovery delay (ϵ). The spectral widths SW_H and SW_N depend on the gyromagnetic ratio of ^1H, the sample length L, and the maximum t_1 evolution time t_1^{max}. In the absence of ^{15}N decoupling, δN and ϵ are both equal to 0 and T_a^* becomes equal to T_a. If additional ^{15}N decoupling is required, T_a is reduced by an amount $\delta N + \epsilon$ in order to keep T_a^* fixed thus covering the same spectral width $SW_H(F_2)$. As a consequence, the strength of the acquisition gradient has to be increased to preserve also the spectral width in F_1. Since the filter bandwidth parameter in ultrafast experiments is directly related to the acquisition gradient strength G_a according to $fb = \gamma_H G_a L$, it follows that for increasing G_a the noise level detected by the receiver increases. From these considerations the relative SNR in STS and standard ultraSOFAST-HMQC spectra is expected to follow the relation

$$\lambda = \frac{SNR(STS)}{SNR(standard)} = \frac{1}{\sqrt{1 - 2(\delta N + \epsilon)SW_H}} \tag{3.14}$$

For a typical amide ^1H spectral width (SW_H) of 4 ppm, a ^{15}N 180° pulse width of 60 µs, and a recovery delay ϵ of 40 mus, we calculate expected SNR gains of $\lambda = 1.4$ at 600 MHz, and lambda $= 1.7$ at 800 MHz. In practice, to reach acceptable acquisition gradient levels, a reduced spectral width and shorter recovery delays need to be chosen. The first results in spectral aliasing in the ^1H (F_2) dimension, while the latter yields residual line splitting (see above) due to imperfect decoupling performance. In the spectra of Fig. 17 an average SNR ratio of $\lambda = 1.8$ is measured with and without the STS pulse element, in close agreement to the expected gain at 800 MHz. At even higher magnetic field strengths, e.g., 900 MHz the sum of ^{15}N pulse length and recovery delay may exceed the required dwell time in the amide ^1H dimension ($T_a = 1/SW_H - 2(\delta N + \epsilon < 0$). Therefore longer T_a^* times would have to be chosen resulting in even more dramatic spectral aliasing. In addition, upon using higher magnetic field strengths stronger ^{15}N decoupling fields are required for the standard version, resulting in partial probe detuning, and consequently in a further decrease of the experimental sensitivity. Therefore STS-ultraSOFAST-HMQC represents a clear improvement in terms of SNR especially for application at high magnetic fields that provide the sensitivity required for ultrafast protein NMR.

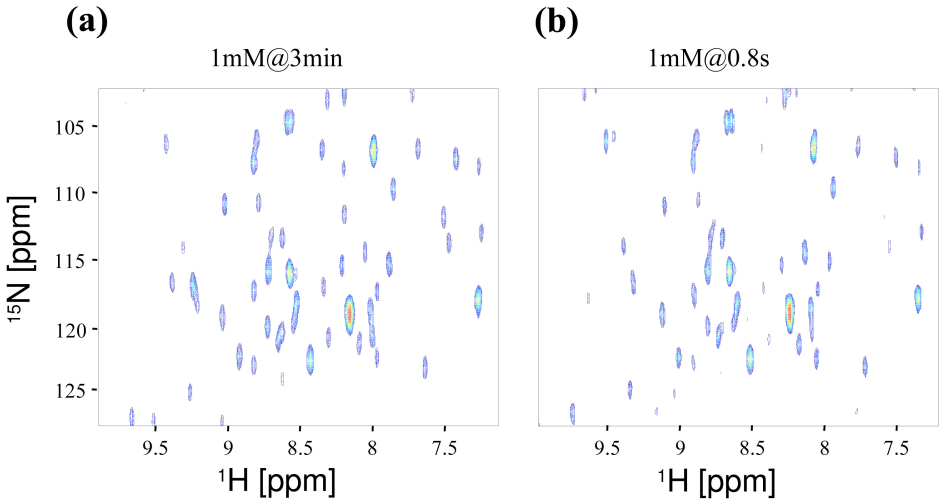


Figure 18 – STS-UltraSOFAST HMQC spectra acquired on a 800 MHz machine equipped with a cryogenic probe. The spectrum on the right has been recorded in a an experimental time of less than 1 second, while a longer acquisition time has been chosen for the left spectrum shown as a reference. Interleaving of 2 acquisitions (scans) with interlaced FT along t_2 was used [120]. Experimental parameters were: $G_e = 36$ G/cm, $N_2 = 128$, $t_1^{max} = 12$ ms, $T_a^* = 250$Êµs, ^{15}N 180° pulse = 62 µs, $G_a = 12$ G/cm. The sample concentration and acquisition time are as indicated on top of each panel.

As demonstrated in Fig. 18, the increased sensitivity provided by the STS-SOFAST-HMQC experiment allows recording a full ^{1}H-^{15}N correlation spectrum of a moderately concentrated (1 mM) ^{15}N-labeled ubiquitin sample in less than 1 s acquisition time on a 800 MHz NMR spectrometer equipped with a cryogenic probe. The spectra were processed using an interlaced FT along t_2, allowing further reduction of the G_a acquisition gradient strength [120]. This experiment presents a promising step towards ultrafast real-time 2D NMR measurements of protein samples at millimolar concentration.

3.3.2 *STS-SOFAST-HMQC*

The STS sequence block of Fig. 16c can also be inserted into a standard SOFAST-HMQC experiment. Although this is not the main focus of the present report, we have set up an optimized pulse sequence for standard time incremented single-transition-spin-state SOFAST-HMQC shown in Fig. 16d. This experiment is conceptually similar to the recently introduced sensitivity-enhanced (SE) IPAP- SOFAST-HMQC [71] where two data sets

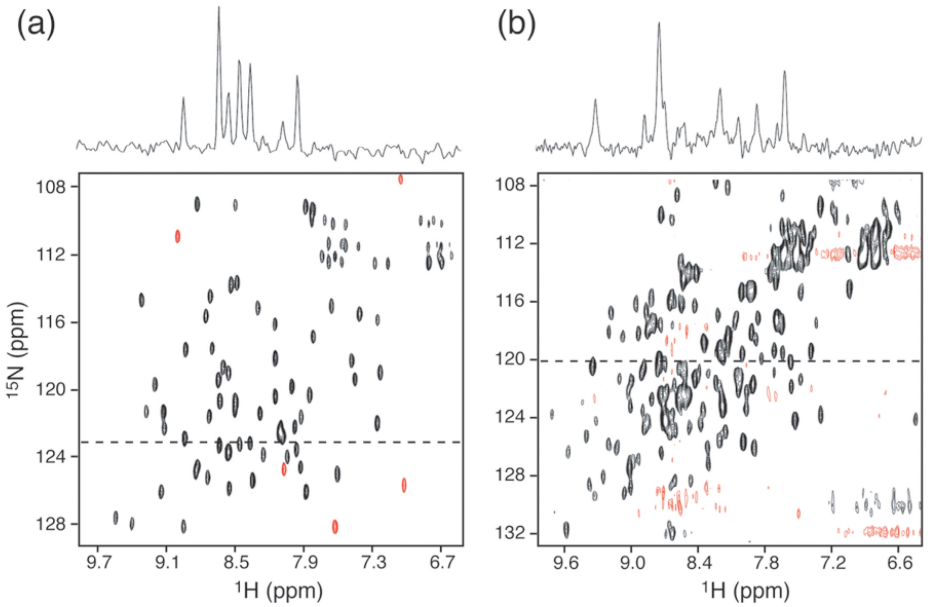


Figure 19 – STS-SOFAST-HMQC spectra recorded on a Varian Inova 800 MHz spectrometer equipped with a cryogenic probe for 0.9 mM samples of (a) ubiquitin (8.6 kDa, 25°C) , and (b) the lipoprotein YajG (20 kDa, 37°C) in experimental times of 4Ês and 8Ês, respectively. Both spectra were recorded with a single scan per t_1 increment, thus without any phase cycling.

are recorded and combined to yield a ^{1}H-^{15}N HMQC correlation spectrum without the need of ^{15}N decoupling during detection. Both, the SE-IPAP and STS versions of SOFAST-HMQC benefit from sensitivity enhancement in the sense that both quadrature components present after the t_1 evolution period contribute to the detected NMR signal (see Eqs. 3.9- 3.12). The STS-SOFAST-HMQC sequence has the advantage that it does not require combination of 2 different data sets, as a single-transition spin state is created by the action of pulsed field gradients. The drawback of this new experiment is that, because of the gradient-induced phase modulation in t_1, the SNR is reduced by a factor $\sqrt{2}$ with respect to either standard SOFAST-HMQC or sensitivity-enhanced IPAP-SOFAST-HMQC. Still the sensitivity provided by this experiment is sufficient to record 2D correlation spectra of proteins at millimolar concentrations in less than 10 s acquisition time, as demonstrated in Fig. 19 for the 2 globular proteins ubiquitin (8.6 kDa) and the lipoprotein YajG (21 kDa). As can be appreciated from the spectra in Fig. 19, the STS-SOFAST-HMQC experiment yields excellent water suppression in a single scan (per t_1 increment) without any phase cycling. The experiment shows also good performance with respect to the time required to reach steady-

state conditions after the creation of fluid turbulences when a fast mixing device is used to initiate a kinetic event inside the NMR magnet. Because the delay between defocusing and refocusing gradient pulses is much shorter in STS-SOFAST-HMQC than in standard SOFAST-HMQC, the performance of STS-SOFAST-HMQC is superior to SOFAST-HMQC and similar to the recently proposed fluid-turbulence-adapted (FTA) SOFAST-HMQC sequence [147].

3.3.3 *Fast amide exchange kinetics monitored by ultrafast real-time 2D NMR spectroscopy*

To experimentally demonstrate the potential of STS-ultraSOFAST-HMQC for monitoring dynamic events in proteins on a seconds time scale, we have measured fast amide deuterium/hydrogen (D/H) exchange rates in ubiquitin at high pH. After rapidly changing the ^{2}H:^{1}H ratio in the sample's buffer, intensity changes of individual cross peaks were monitored in a series of real time 2D spectra to yield a quantitative measure of D/H exchange rates [13, 147]. Such amide hydrogen exchange measurements provide atom-resolved information about the (partially) unfolded conformational states of a protein that are present under native conditions, even if they are only populated to minuscule levels. The ability to detect small amounts of unfolded conformations arises from the fact that amide hydrogen atoms are protected from exchange as long as they are hydrogen-bonded or buried within the protein. Therefore, only when the amide hydrogen gets solvent exposed by some local or global unfolding event, it can be replaced by a proton or deuteron from the solvent. One interesting case of amide hydrogen exchange arises when the measured exchange rates can give direct insight into the kinetics of protein unfolding. This accounts if the intrinsic chemical exchange rate is fast compared to the closing (refolding) rate from an exchange-competent to an exchange-protected conformation. Under these conditions, generally reached at high pH values, exchange is limited by the rate at which an amide site gets exchange-competent, and the NMR-observed exchange rates will thus directly reflect the unfolding rates of the individual residues [19].

For the D/H exchange experiment shown in Fig. 20, ^{15}N-labeled ubiquitin was dissolved in 90 µl D_2O buffer at pH 12 at a concentration of ≈8 mM. The hydrogen exchange reaction was active for a time long enough to ensure complete deuteration of all amide sites in the protein. This solution was then injected into an NMR sample tube containing 350 µl of H_2O buffer that

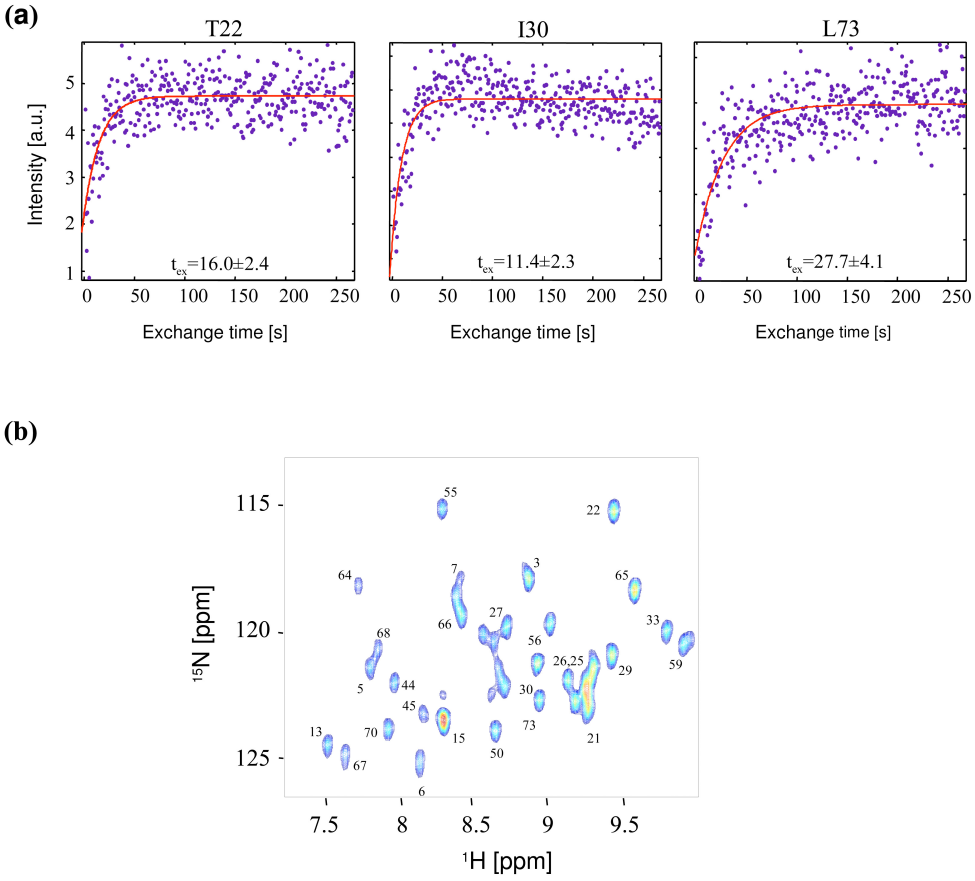


Figure 20 – H/D exchange rates measured on a 600 MHz spectrometer equipped with a cryogenic probe. Uniformly ^{15}N-labeled ubiquitin in D_2O-buffer was concentrated to a volume of 90µL and loaded into a fast-injection device. H/D exchange was initiated by injecting the ubiquitin-D_2O-buffer solution into 350 µL of H_2O buffer inside the spectrometer. The dead time between mixing and beginning of data acquisition was less than one second. A series of single-scan STS-UltraSOFAST HMQC experiments of 350 ms duration per 2D spectrum was recorded to follow the H/D exchange kinetics. Phase cycling was achieved in consecutive 2D acquisitions, as described previously [147], that does not alter the time resolution of the kinetic dimension. Experimental points in (A) reflect peak heights extracted from the STS-UltraSOFAST HMQC 2D spectra; the t_{ex} exchange lifetimes given in the figure were obtained by fitting of the data points to the equation $I(\tau) = I_0 + I_{inf}(1 - exp\left(-\tau/t_{ex}\right)$. In (B) is shown a reference spectrum acquired as soon as the experiment was finished with the same experimental parameters but with 128 scans. The numbers near the peaks correspond to the numbers above each exchange curve.

had been placed inside the NMR magnet. The dead time between injection and data acquisition was less than 1 s. After injection the exchange process was followed by recording a series of STS-ultraSOFAST-HMQC spectra with a repetition rate of $\approx$3 s^{-1} (350 ms per 2D spectrum). The peak intensity as a function of exchange time is plotted in Fig. 20a for three distinct amide sites, corresponding to residues T22, I30, and L73. Figure 20b shows the final spectrum after completion of the exchange process (to increase the SNR 128 spectra were added for this spectrum). The peak intensity increases as more and more deuterons get replaced by solvent protons, reaching a plateau value corresponding to the H_2O/D_2O ratio in solution (in our experiment 3.9:1). This means that although the protein concentration inside the NMR tube was $\approx$1.6 mM, the hydrogenated, and detectable amide sites corresponded to an effective protein concentration of only 1.27 mM.

As the buildup curves in Fig. 20a indicate, the chosen acquisition time of 350 ms per spectrum is a hard test case for the technique in terms of sensitivity and measurement accuracy. Exchange time constants of a second or below can not yet reliably been measured using this technique, although the repetition rate of a few Hz in principle provides access to this time scale. Therefore, the real-time 2D NMR data shown in Fig. 20 should be taken as a proof of principle. It can be foreseen that experimental sensitivity will continue to increase over the years, and we expect that with this, real-time studies of molecular kinetics at multi-Hz rates will become of practical use. In any case, the STS-ultraSOFAST HMQC experiment presents a step forward towards this aim. Furthermore, and in contrast to previously proposed schemes for fast site-resolved kinetics studies, ultraSOFAST NMR does not require any prior knowledge about peak positions as required for Hadamard spectroscopy [13, 86] or extensive spectral aliasing strategies [95].

3.4 CONCLUSIONS

We have introduced STS-SOFAST-HMQC, a sequence that allows the acquisition of 2D ^{1}H-^{15}N correlation experiments yielding a single cross peak per residue within one scan without heteronuclear decoupling and without the need to record separate in-phase and anti-phase experiments. The full power of this pulse sequence element comes into place when combined with spatial encoding, where the previously applied heteronuclear decoupling using 180° rf pulses leads to significant losses in sensitivity, and decoupling performance is less-than-optimal due to interfering effects from the acquisi-

tion gradients. We have shown that well resolved 2D correlation spectra can be obtained at a repetition rate of several Hertz, further extending the time scale accessible to multidimensional real-time kinetic NMR.

4

GENERALIZATION OF LONGITUDINAL RELAXATION ENHANCEMENT BASED OPTIMIZATION OF NMR EXPERIMENTS

4.1 THE BAND-SELECTIVE EXCITATION SHORT TRANSIENT (BEST) EXPERIMENTS

The pulse sequences presented in the previous chapters demonstrate how a combination of longitudinal relaxation enhancement and Ernst angle excitation can speed up significantly the acquisition of 2D ^{1}H-^{15}N or ^{1}H-^{13}C correlation spectra. While Ernst angle excitation is limited to rather simple pulse sequences with few pulses involved, such as the SOFAST HMQC experiments, longitudinal relaxation enhancement can still provide a significant sensitivity and speed gain in more complex pulse sequences. Thus, this approach can be extended to even higher dimensional experiments, with a resulting increase in signal to noise for high repetition rates.

One example is the family of Band-selective Excitation Short Transient (BEST) experiments. A whole set of triple resonance experiments needed for resonance assignment of proteins has been proposed recently [94]. Similar to the SOFAST experiments selective ^{1}H pulses are implemented to allow minimal perturbation of proton spins resonating outside of the excitation band, permitting enhanced longitudinal relaxation.

Figure 21 shows the example of an BEST HNCA sequence. The coherence transfer pathway is the same as for standard sensitivity enhanced HNCA. The selective pulses are chosen to excite exclusively amide protons. The optimized pulse sequences give a significant sensitivity gain of a factor of around 1.5 [94, 149], and the possibility to speed up the spectra acquisition from at least several hours down to a few minutes. For example, a HNCO experiment can be acquired on a 1 mM Ubiquitin sample in 15 min with

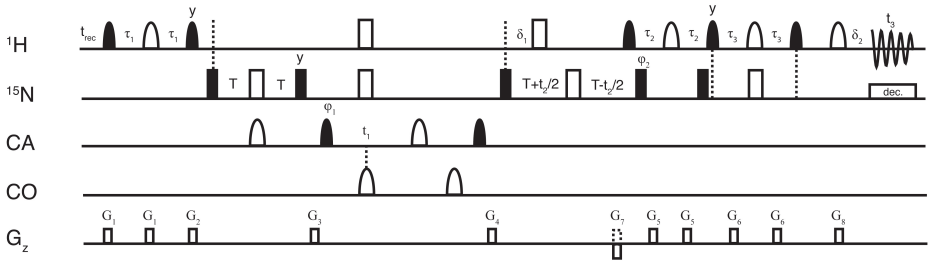


Figure 21 – A BEST HNCA pulse sequence to record HNCA correlation spectra in a few minutes [149]. Filled and open pulses indicate 90° and 180° rf pulses, respectively. Quadrature detection in t_1 and t_2 is obtained by time-proportional phase incrementation $\phi_1 = x, y$ and $\phi_1 = x, -x$. The sign of G_7 is inverted for experiments with $\phi_2 = -x$. The delay δ can be adjusted to suppress J_{NH} coupling.

uniform time-domain sampling [94, 149]. As the set of standard assignment experiments includes generally several experiments, this approach can reduce spectrometer-time needed for protein backbone assignment dramatically. Additionally, the experiments can easily be combined with sparse sampling techniques to further reduce the experimental time. This spectroscopic "trick" is theoretically extendable to a variety of NMR experiments based on an initial magnetization transfer from a subgroup of ^{1}H spins to other heteronuclear spins.

Generally, the selection of a proton spin subset is realized by the use of "band-selective" rf pulses that have been optimized to yield a "top hat" response in frequency (chemical shift offset) space that is uniform (constant rotation angle) over the chosen band width (excitation band), and close to zero (no effect) for spins resonating outside this spectral window (suppression band). A narrow transition region exists where spin evolution is undefined. Many efforts have been made in the early 90s in developing such band-selective pulses using numerical optimization methods, resulting in a large variety of different pulse shapes. For the example of the BEST sequences, often a combination of BURP pulses [42], PC9 pulses [127], and SNOB pulses are used for selective excitation [81]. Other popular band-selective pulses used today for biomolecular applications are the Gaussian pulse cascades of Emsley and Bodenhausen [27, 28].

Unitary spin transformations in weakly coupled spin systems can be described by a series of three-dimensional rotations in product operator space [161]. These simple concepts that form the basis of hundreds of NMR experiments designed for various chemical and biochemical applications,

only apply as long as chemical shift and spin coupling evolution during radio-frequency (rf) pulses can be safely neglected. This is generally the case for "hard" pulses, characterized by a constant high rf power, and a short overall duration of a few microseconds. The situation is more complicated for shaped rf pulses optimized for frequency-selective, or broad-band spin manipulation that often require much longer pulse durations resulting in non negligible chemical shift and J-coupling evolution during the shaped pulse. An additional problem associated with shaped pulses is that they have been optimized for a particular purpose (spin rotation) such as spin excitation ($I_z \rightarrow I_{xy}$), or spin inversion ($I_z \rightarrow -I_z$), while generally no specification is made for the effect of the pulse on other initial spin states. So called general-rotation pulse shapes that yield similar rotation properties in spin space than hard pulses have also been designed for certain purposes. In practice, however, they are often more sensitive (than single-rotation pulses) to pulse imperfections, such as B_1 field inhomogeneity or rf mis-calibration, and spin-coupling evolution during the pulse has not taken into account for the pulse optimization. For these reasons it may be a non-trivial task to replace hard pulses by shaped pulses in common pulse sequences to make them either band-selective or broadband. The following section will address this problem.

4.2 Guidelines for the use of band-selective radiofrequency pulses in heteronuclear NMR: example of longitudinal-relaxation-enhanced BEST-type ^{1}H-^{15}N correlation experiments

In this section we demonstrate that for certain pulse shapes their effect on a scalar-coupled heteronuclear two-spin system (neglecting spin relaxation effects) can be approximated by one (or two) delay(s), accounting for possible chemical shift and coupling evolution during the pulse, and an "ideal" rf pulse rotation. Using these "binary" (delay/pulse) representations it becomes straightforward to replace standard hard rf pulses in complex multi-pulse sequences by the appropriate pulse shape, and to properly adjust transfer and chemical-shift editing delays to account for spin evolution during the shaped pulses, even in cases where several coherence-transfer pathways need be realized at the same time. We will discuss the use of shaped pulses in the context of longitudinal-relaxation-enhanced 2D BEST-HSQC [94, 149], and 2D BEST-TROSY [31] pulse schemes, recently developed in our laboratory, and the latter presented here for the first time for protein applications. Although there exists already an abundant literature on selective pulses,

revisiting this topic in the context of recently developed biomolecular NMR experiments will be of practical use for many NMR spectroscopists interested in deriving new pulse schemes, or optimizing well-established correlation experiments by replacing hard pulses with the appropriate band-selective pulse shapes.

To gain a more detailed understanding of spin evolution during the selective pulses used in BEST-type experiments, we performed simulations for a scalar-coupled ^{1}H-^{15}N spin system. We have used the SPINEVOLUTION software [172] to calculate the spin evolution during various pulse shapes. All simulations have been carried out for a ^{1}H Larmor frequency of 600 MHz, a heteronuclear coupling constant of 100 Hz, and neglecting spin relaxation effects. The effect of a selective pulse on each of the 15 (non-unity) spins states of the density matrix was computed as a function of ^{1}H frequency offset varying from -5 to +5 kHz. ^{15}N offset effects were neglected as the radiofrequency field in the kHz range applied at the Larmor frequency of ^{1}H does not directly affect the ^{15}N spins. Figure 22 shows the simulation results obtained for an EBURP-2 pulse of duration σE=1.92ms, yielding spin excitation over a bandwidth of about 4 ppm. A total of 40 representative examples, out of the 225 (15 x 15) possible coherence transfer pathways are shown. EBURP-2 has been numerically optimized for performing a $H_z \rightarrow H_y$ rotation over the excitation bandwidth, but interestingly also other coherence transfer pathways are quite efficient. For example, if applied to transverse ^{1}H coherence H_x or H_y part of the coherence is transferred to H_z spin polarization with a linear dependence on the ^{1}H frequency offset. Also some scalar-coupling mediated pathways are observed, e. g. $H_x \rightarrow 2H_xN_z$, $2H_xN_z \rightarrow H_x$, $N_x \rightarrow 2H_yN_y$, and $2H_zN_x \rightarrow N_y$. The observed linear ^{1}H offset-dependence motivated us to search for a suite of "ideal" 90° pulses and free evolution delays that is able to reproduce the coherence transfer amplitudes achieved by the EBURP-2 pulse within the excitation band (fig. 22b). Such a binary sequence, comprising one delay (white box) accounting for ^{1}H chemical shift and heteronuclear scalar coupling evolution, and one general 90° rotation pulse (black box), is shown in fig. 23a (center panel) for EBURP-2. The length of the black box, during which ^{1}H chemical shift and heteronuclear scalar coupling evolution is inactive, has been adjusted to correctly account for ^{15}N chemical shift evolution during the shaped pulse duration. For computational purpose, the white box is replaced by a delay of length $\lambda\sigma E$, while the black box is approximated by an instantaneous 90° pulse of the same phase as the shape. The duration of the black box can be neglected as no ^{15}N chemical shift evolution is taken into account

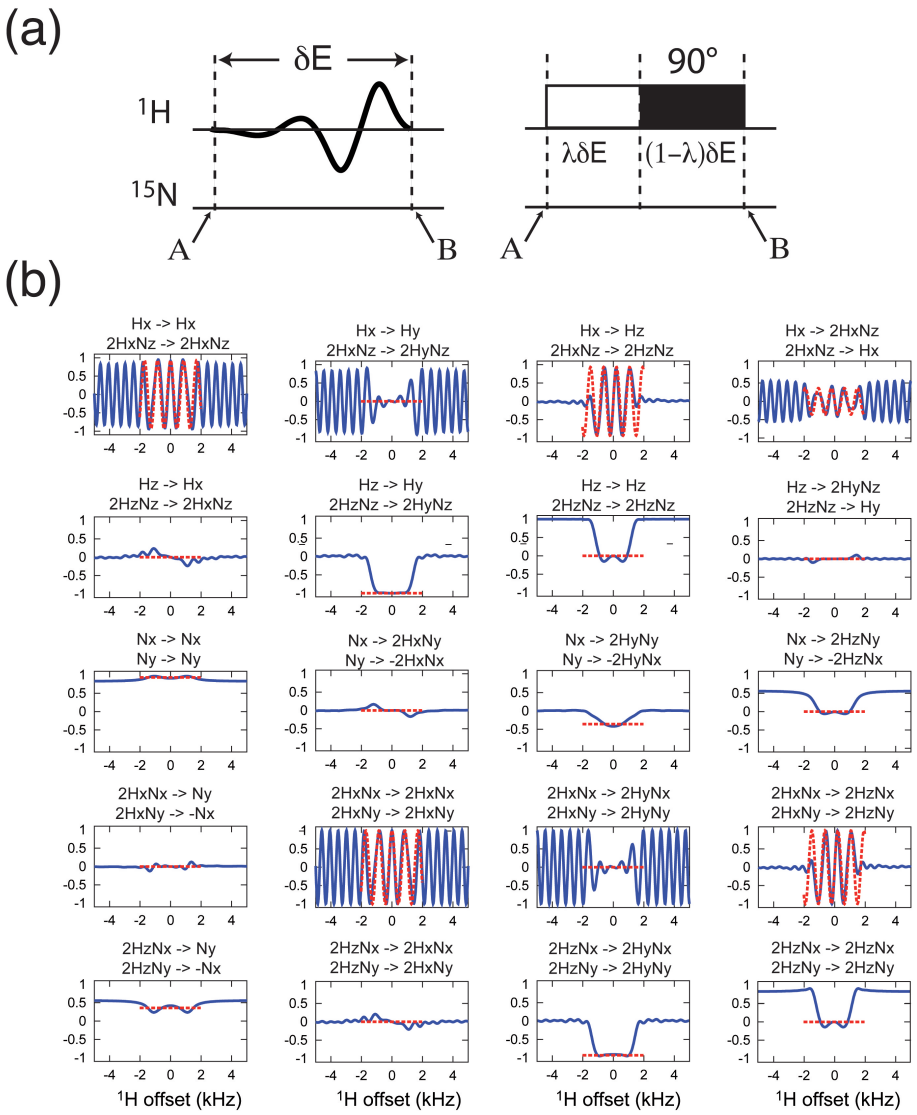


Figure 22 – Numerical simulations of the effect of EBURP-2 [42] on various initial spin states of a scalar-coupled ^{1}H-^{15}N spin system ($J_{HN} = 100$Hz). (a) Calculations were performed for the real pulse shape (left panel), and for a binary (rf-on/off) equivalent (right panel). The EBURP-2 pulse length was set to 1.92 ms corresponding to an excitation bandwidth of about $\pm$1.2kHz. (b) Selection of representative coherence transfer pathways computed for EBURP-2 (blue lines) and the corresponding binary scheme (red dashed lines).

for this simulation. The transfer amplitudes computed from this binary scheme are shown as a function of frequency offset in the graphs of fig. 22b for different pathways (red dashed lines), and superposed on the results obtained for the EBURP-2 pulse (blue lines). The λ value has been optimized by minimizing the discrepancy between the transfer amplitudes calculated for the real pulse shape and the binary scheme in the offset range ± 1.2 kHz. $\lambda = 0.62$ means that ^{1}H chemical shift and ^{1}H-^{15}N scalar-coupling evolution is active for a fraction 0.62 of the total shaped pulse duration. Overall, an excellent agreement is obtained when replacing the EBURP-2 pulse by its binary representation. The deviation is less than 10% for all coherence transfer pathways and all offset frequencies within the excitation band. Therefore such optimized bimodal schemes present a convenient tool for using shaped pulses for various purposes, others than for what they have been initially conceived, in complex multi-pulse sequences. Similar calculations as those shown for EBURP-2 in fig. 22 were done for other shaped pulses commonly used in biomolecular NMR spectroscopy. Figure 23 shows the amplitude-modulated pulse shapes, the numerically optimized binary schemes, and the excitation (or inversion) profiles obtained for (a) the excitation (E) pulses EBURP-2 [42] and E400B [172], Q5 [28], PC9 [83] (b) the inversion (I) pulse ISNOB-5 [81], and the refocusing (R) pulses REBURP [42] and Q3 [28]. General 90° rotation pulses, such as UBURP [42] were not considered here as these particular pulse shapes are known to be particularly sensitive to B_1-field inhomogeneity and B_1-field calibration. Time reversal of the pulse shape obviously also translates to a time reversal of the binary scheme shown in fig. 23a. As a consequence, the only difference between an excitation pulse and a time-reversed excitation pulse, further on called a flip-back (FB) pulse, is whether chemical shift and spin coupling evolution takes place before or after the general 90° spin rotation. Slight changes in the amplitude modulation profile may translate into a significant effect on the binary scheme, as illustrated for the excitation pulses EBURP-2 and E400B that have very similar shapes. Finally, our numerical simulations show that pulse shapes optimized for the same purpose, e.g. the refocusing pulses REBURP and Q3 may show a completely different behavior with respect to chemical shift and coupling evolution during the pulse. While during the symmetric REBURP shape, spin evolution occurs to good approximation during the entire pulse length with an instantaneous spin inversion in the centre of the pulse, the asymmetric Q3 pulse shape is best described by a single 180° rotation without any evolution delays. Of course the net effect on chemical shift and spin coupling evolution is the same for the 2 pulse shapes. The difference comes into play when an additional 180° ^{15}N pulse

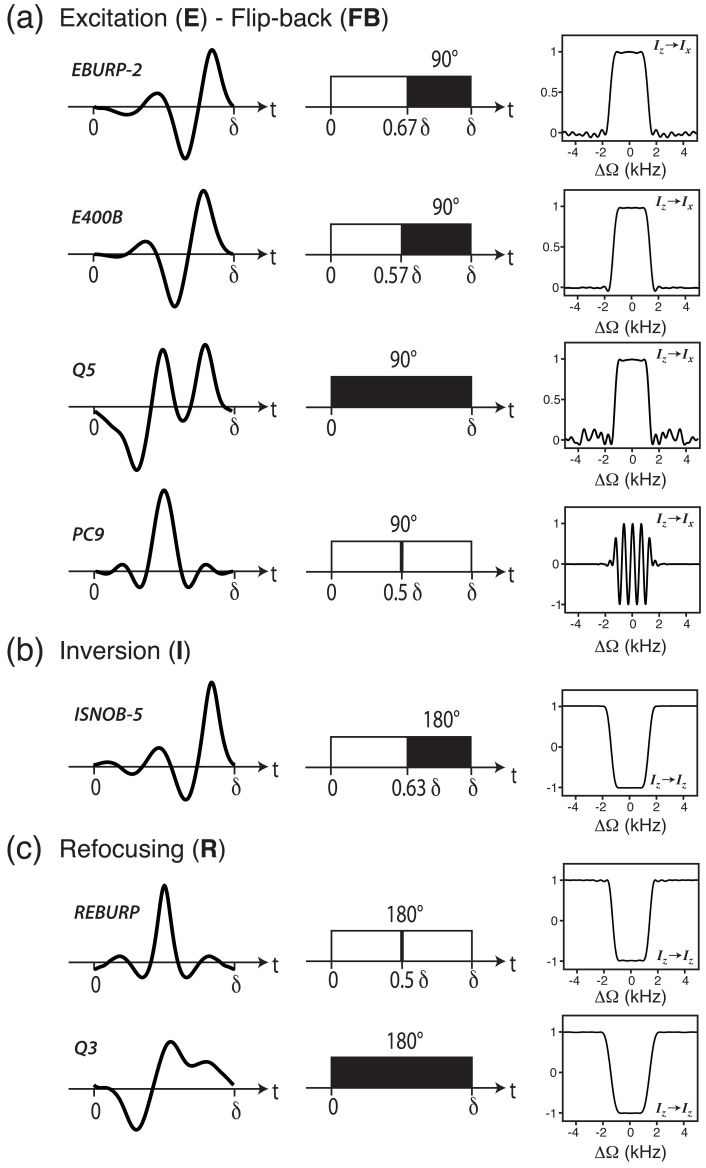


Figure 23 – Amplitude modulation functions, binary schemes optimized by numerical simulations, and excitation (inversion) profiles for several 90° (180°) pulses proposed in the literature. The pulse length have been adjusted for an excitation/inversion bandwidth of ±1.2kHz. (a) Excitation pulse shapes EBURP-2 [42], E400B [172], Q5 [28], and PC9 [83]; (b) the inversion pulse ISNOB-5 [81]; the refocusing pulses REBURP [42] and Q3 [28]. Note that the ISNOB-5 pulse shape is very similar to IBURP [42].

is applied during the ^{1}H shaped pulse in order to introduce heteronuclear coupling evolution (see fig. 24).

With the bimodal schemes of figure 23 in hand, it now becomes straightforward to replace hard rf pulses in common pulse sequence elements by the appropriate band-selective pulse shapes, as we will demonstrate in the remaining of this article. As a first example, different experimental realizations of a ^{1}H band-selective heteronuclear INEPT transfer are shown in figures 24a-d. As demonstrated by numerical simulations, all of these INEPT schemes yield a similar performance with respect to the resulting transfer amplitudes as long as the transfer delays are properly adjusted in order to take into account spin coupling evolution during the different pulse shapes as given by the corresponding replacement scheme. In particular, it does not matter whether spin excitation is realized by an excitation or flip-back pulse shape (figs. 24a and 24b), or whether an inversion or refocusing pulse shape is used for general 180◦ rotation purposes (fig. 24d). Let us now discuss more complex pulse schemes where two or more coherence-transfer pathways have to be realized at the same time. Figure 25 shows the experimental realization of a sensitivity-enhanced $^{15}N \rightarrow ^1$ H transfer element required for the preservation of equivalent pathways (PEP) [83], and implemented for echo/antiecho-type ^{15}N quadrature detection in ^{1}H-^{15}N HSQC-type experiments [127]. This sequence, referred to in the following as planar mixing scheme [68], converts the 2 orthogonal antiphase coherences $2N_xH_z$ and $2N_yH_z$ present after the ^{15}N chemical shift editing (t_1) period, into detectable ^{1}H inphase coherence H_x and H_y respectively. As a consequence of the simultaneous occurrence of the 2 coherence-transfer pathways, each of the 90◦ ^{1}H pulses (except for the first one) needs to perform 90° spin rotations for 2 orthogonal spin components, and heteronuclear spin coupling evolution during the pulses can not be neglected. The pulse sequence of fig. 25b shows one possible solution to this problem that is based on the use of excitation, flip-back, and refocusing pulse shapes only, and does not require any general 90° rotation pulse. Numerical simulations, shown in fig. 25c, illustrate the good performance of this pulse scheme with respect to the transfer amplitudes obtained for the 2 coherence transfer pathways. It should be pointed out that ^{15}N chemical shift evolves during the whole duration (σE) of the first excitation pulse (E_y), while ^{1}H-^{15}N scalar coupling evolution only occurs during a fraction σE (with $\lambda = 0.62$ in the case of EBURP-2). Therefore additional 180° ^{15}N and ^{1}H pulses are applied after the t_1 period with the delays between them properly adjusted to refocus the two effects (fig. 25b). A band-selective double-S^{3}CT sequence [134, 153]

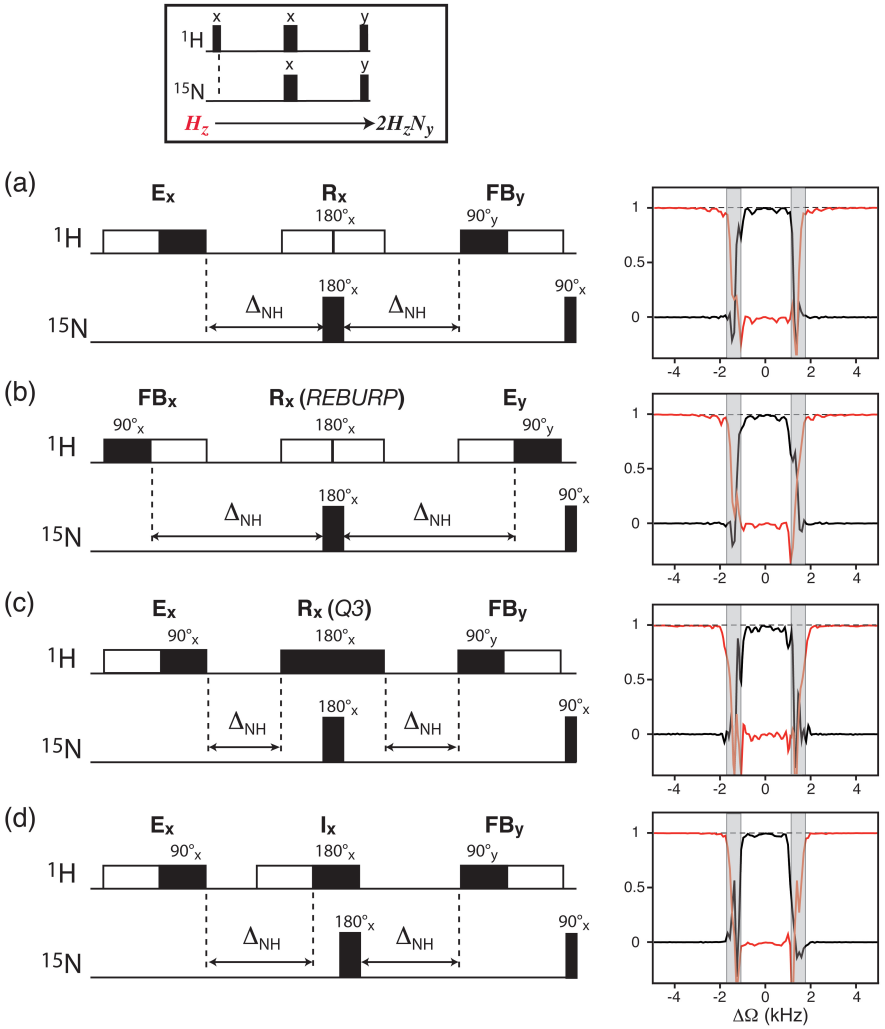


Figure 24 – Experimental realizations of a heteronuclear ^{1}H-^{15}N INEPT transfer using different band-selective ^{1}H pulse shapes, and transfer delays adjusted according to the bimodal schemes introduced in figure 22. For all sequences, the amplitudes of the initial (red: Hz) and final spin states (black: $2H_zN_y$), computed as a function of ^{1}H resonance offset are shown on the right side.

that performs spin-state-selective $^{15}N \rightarrow ^{1}H$ transfer is shown in figure 26a. Again, 2 coherence-transfer pathways need to be realized to convert a ^{15}N single-transition state into a ^{1}H single-transition state. Transfer amplitudes close to 1 are achieved for both pathways (see fig. 26c) when correctly taking into account spin evolution during the shaped pulses as easily done on the basis of the binary representations presented in figure 23. The band-selective INEPT, planar mixing, and double-S^3CT pulse schemes presented in figs. 24, 25, and 26, respectively, form the basis of the recently introduced sensitivity-enhanced BEST-HSQC [94, 117] and BEST-TROSY [31] experiments that offer new NMR tools for fast and sensitive collection of protein and nucleic acids spectra. In our original work, the various transfer delays were adjusted empirically for optimal sensitivity of the pulse schemes. The resulting values were close, but not identical to the theoretically determined optimal values reported here. Therefore, although we could demonstrate the good experimental performance of these sequences, a detailed understanding of the spin evolution during the shaped pulses used in these sequences was still missing. The bimodal schemes, obtained from numerical simulations and presented here for various band-selective pulse shapes can now be used for choosing the appropriate pulse shape for a given situation, and for adjusting pulse sequence delays without the need for experimental optimization. Figure 27 shows an experimental characterization and comparison of ^{1}H-^{15}N sensitivity-enhanced BEST-HSQC and BEST-TROSY correlation experiments. NMR data were recorded for different ^{15}N-labelled, and fully protonated protein samples, ubiquitin (8.6 kDa, 2 mM), β2-microglobulin (12 kDa, 0.3 mM), and YajG (21 kDa, 0.2 mM), on a 600 MHz Varian spectrometer equipped with a cryogenically cooled triple-resonance probe and pulsed z-field gradients. EBURP-2, time-reversed EBURP-2 and REBURP pulses were used for excitation (E), flip-back (FB), and refocusing (R) purposes, as in our hands this combination gave the best results. The experimental sensitivity, defined as the spectral intensity normalized for a given experimental time and equal noise level, has been determined experimentally as a function of the scan time, and is plotted in fig. 27a for BEST-HSQC, and in fig. 27b for BEST-TROSY. The scan time comprises the pulse sequence duration including the detection period, as well as the recycle delay. A similar behavior is found for all 3 proteins with an optimal scan time of 0.5 s for BEST-HSQC, and a slightly reduced value of 0.4 s for BEST-TROSY. These optimal scan times, for which the pulse sequences yield highest sensitivity, seem to be independent of the protein size, indicating that the increased ^{1}H spin-diffusion efficiency in larger proteins with increased tumbling correlation times does not signif-

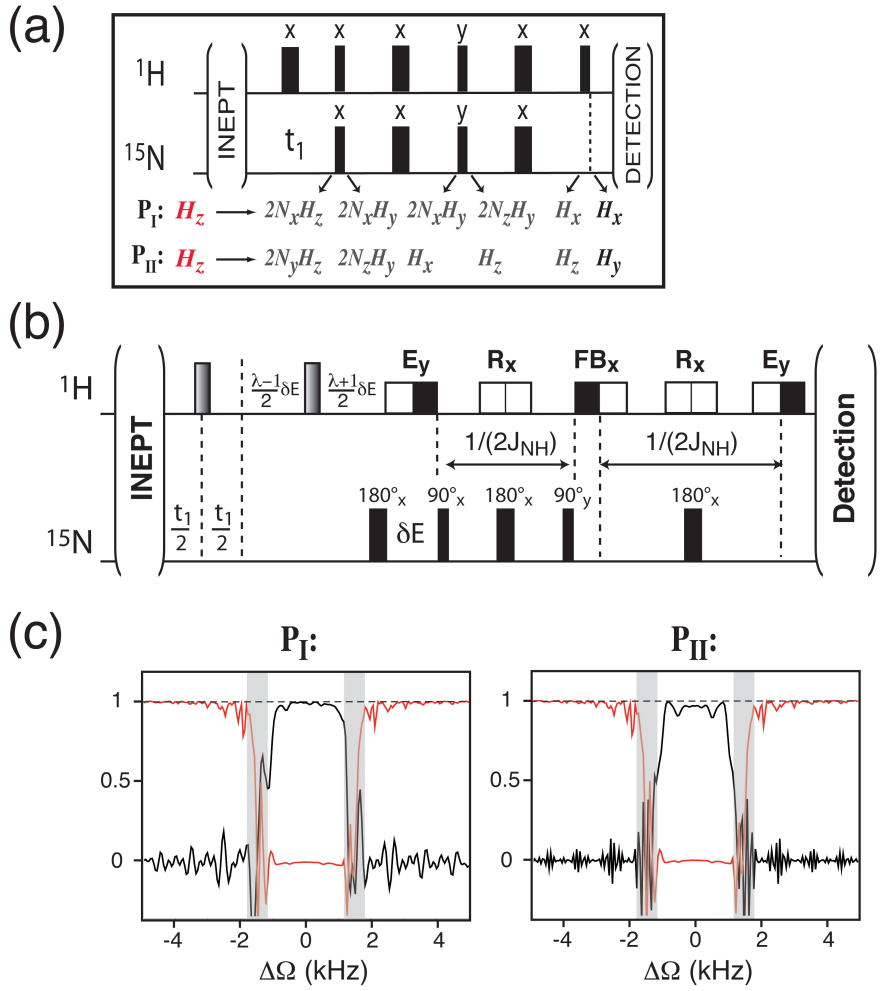


Figure 25 – (a) Planar mixing sequence as required for sensitivity-enhanced ^{15}N-^{1}H back-transfer in HSQC-type correlation experiments. (b) Band-selective version based on excitation, flip-back and refocusing pulse shapes, and optimized transfer delays. The notation σE stands for the length of the ^{1}H excitation pulse shape (E), while the scaling factor λ corresponds to the percentage of the pulse duration that needs to be taken into account for spin coupling and chemical shift evolution according to the schemes of figure 1. The ^{1}H 180° pulses applied during the t_1 period are preferably applied with a shape optimized for broadband spin inversion [72] (c) Amplitude profiles of the initial (red lines) and final (black lines) spin states of the 2 coherence transfer pathways computed as a function of ^{1}H resonance offset.

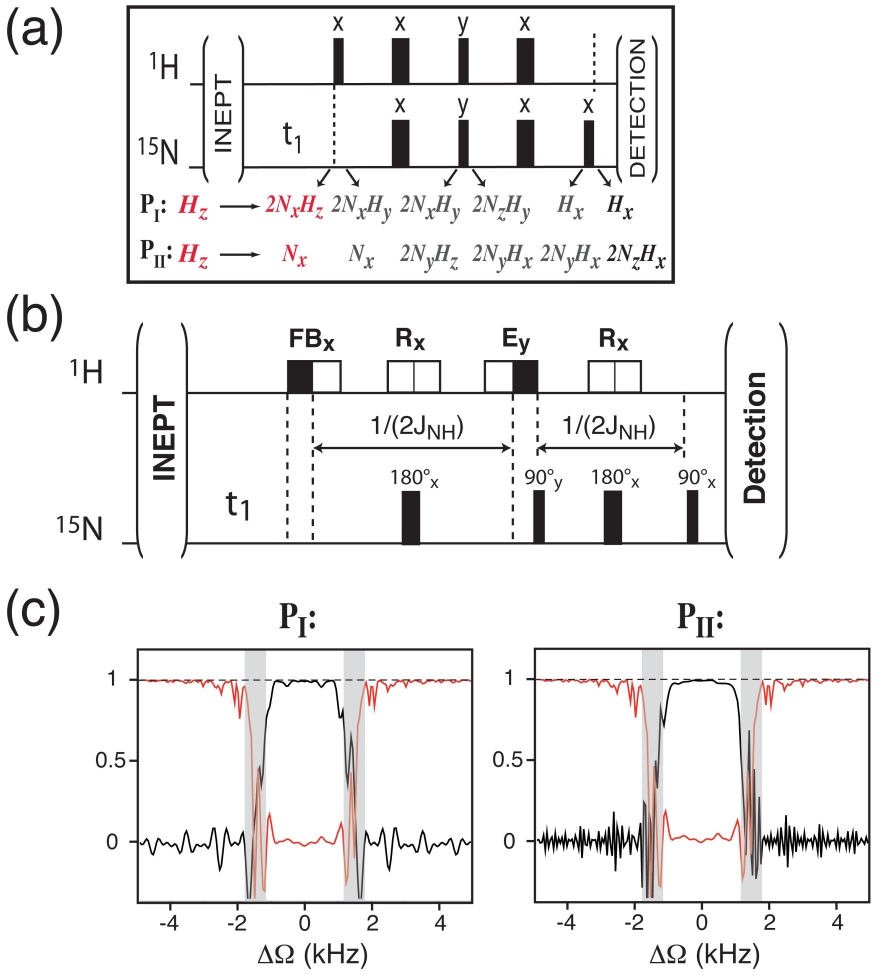


Figure 26 – (a) Double-S^{3}CT sequence as required for single-transition-state-selective ^{15}N-^{1}H back-transfer in TROSY-type correlation experiments. (b) Band-selective version based on excitation, flip-back and refocusing pulse shapes, and optimized transfer delays. (c) Amplitude profiles of the initial (red lines) and final (black lines) spin states of the 2 coherence transfer pathways computed as a function of ^{1}H resonance offset.

icantly further reduce the selective T_1 relaxation times of amide protons. The slightly shorter optimal scan time found for BEST-TROSY is most likely explained by the smaller number of ^{1}H pulses required in this sequence yielding less disturbance of non-amide protons, and the absence of ^{15}N decoupling during data acquisition that may result in partial probe detuning at high repetition rates [144]. 2D BEST-HSQC spectra (left panel - black contours) and BEST-TROSY spectra (right panel - red contours) recorded under optimal sensitivity conditions are shown for ubiquitin (fig. 27c), β2-microglobulin (fig. 27d), and YajG (fig. 27e). While for ubiquitin the BEST-HSQC yields a slightly higher sensitivity than the BEST-TROSY, for the other proteins the sensitivity obtained by the two BEST pulse schemes is comparable as illustrated by the 1D traces shown on the left of figs. 27c-e. Because for an amide ^{1}H-^{15}N spin pair the CSA-dipolar cross-correlation becomes more efficient at higher magnetic field strength, with a maximum reached at about 1 - 1.2 GHz ^{1}H frequency, it is expected that at highest currently available magnetic field strength (800 to 1000 MHz) the BEST-TROSY experiment outperforms the BEST-HSQC sequence both in terms of sensitivity and spectral resolution. Typically a 30-100% sensitivity gain is achieved for (protonated) proteins or nucleic acids when using ^{1}H-^{15}N BEST-type experiments instead of standard hard-pulse-based pulse sequences if both are optimized independently for optimal sensitivity [31, 94, 149]. Of course, these BEST sequence blocs (figs. 23-25) are equally useful for higher-dimensional triple-resonance H-N-C type correlation experiments to increase experimental sensitivity or to reduce experimental times.

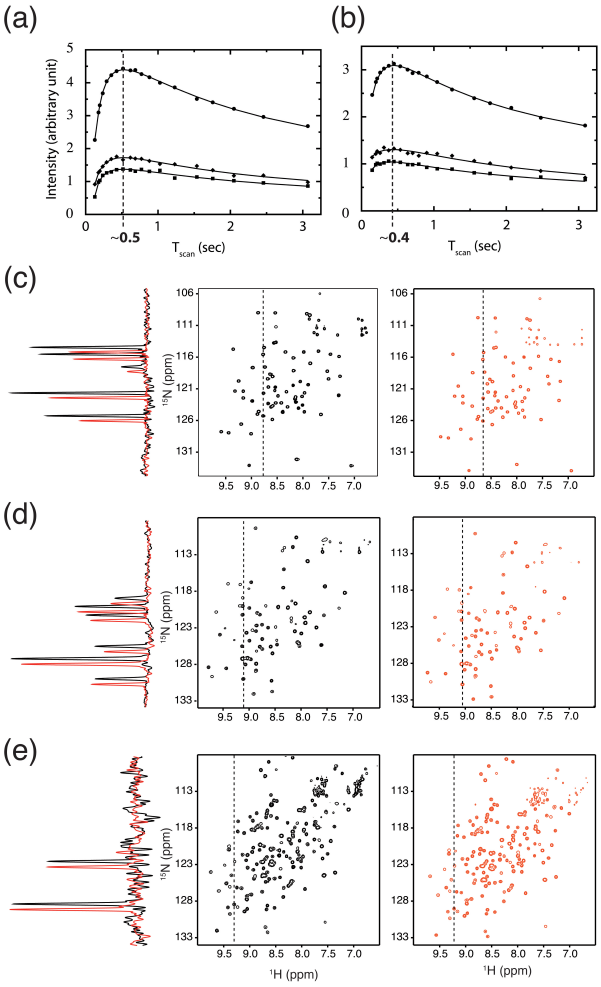


Figure 27 – Experimental data recorded for 3 different proteins on a 600 MHz Varian spectrometer equipped with a triple-resonance cold probe. Sensitivity curves obtained from a series of (a) 1D BEST-HSQC, and (b) 1D BEST-TROSY spectra. For the initial INEPT transfer the pulse sequence of fig. 23a was used, and for the back transfer the sequences of fig. 26 (HSQC) and fig. 26 (TROSY) were used as described. The pulse shapes EBURP-2, time-reversed EBURP-2, and REBURP were applied for excitation (E), flip-back, and refocusing (R), respectively. The integrated spectral intensities, normalized for equal measurement times and scan number are plotted as a function of scan time (pulse sequence duration including acquisition time plus recycle delay) for ubiquitin (circles), β2-microglobulin (diamonds), and YajG (squares). 2D BEST-HSQC (left Ð black contours) and BEST-TROSY (right Ð red contours) are shown in (a) for ubiquitin, (b) for β2-microglobulin, and (c) for YajG. For all experiments the ^{1}H acquisition time was set to 70 ms, and 100 complex t_1 points were recorded for a ^{15}N spectral width of 1800 Hz. The recycle delay was set to 0.43 s for HSQC and 0.33 s for TROSY experiments, and the spectral intensity was normalized for equal experimental times. In addition 1D ^{15}N traces extracted at the ^{1}H frequency indicated by a dashed line in the 2D spectra are shown.

4.3 BEST TYPE OPTIMIZATION OF CHEMICAL EXCHANGE EXPERIMENTS

In the previous chapters the focus has been on optimized pulse sequences for real time NMR studies of biomolecular kinetics. It has been demonstrated that longitudinal relaxation enhanced pulse sequences are very beneficial tools to increase the time resolution of mulitdimensional NMR techniques used in this field. Real time NMR is a powerful tool to study molecular kinetic processes in the order of a few s to min. Many biological relevant processes, as domain motions [67], enzymatic recognition and catalysis [1, 26], or protein unfolding/folding [145, 167] occur on a timescale of µs to rarely several s, and still only very slow processes are accessible to real time studies (figure 28). However, it is often of considerable interest to characterize the dynamics of the conformational transitions that are involved in these processes. One class of experiments that has been shown to be among the most useful are experiments probing the modulation of the isotropic chemical shift by the dynamic process. This mechanism is called chemical (or conformational) exchange and can in the simplest case be described by an equilibrium between two different populated states. The kinetics can then be defined over NMR accessible parameters such as the exchange constants, the populations of the excited states, and the chemical shift difference between both states, containing valuable structural information.

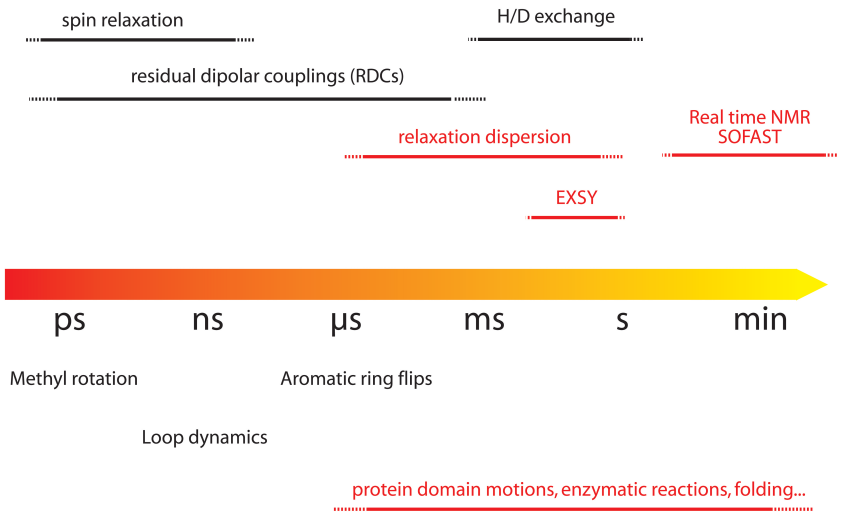


Figure 28 – Schematic representation of the timescales of different biochemical kinetics in high-field NMR and liquid-state NMR techniques used for their characterization.

Exchanges processes can be monitored by longitudinal (exchange spectroscopy) and transversal (relaxation-dispersion) relaxation type experiments.

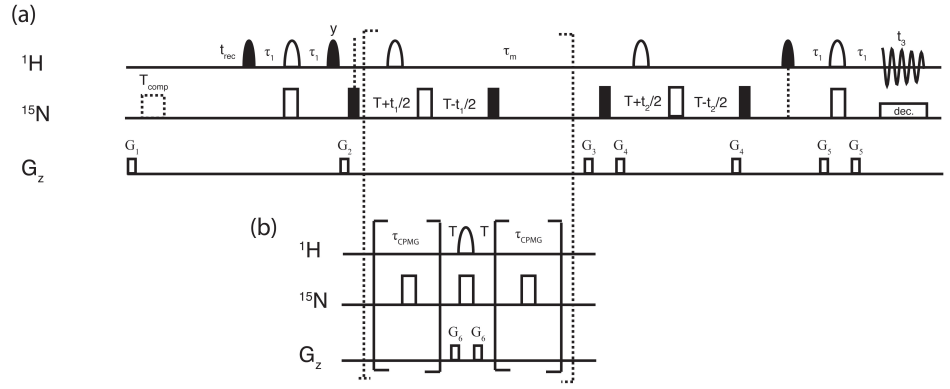


Figure 29 – Pulse sequences for (a) BEST-exchange spectroscopy and (b) BEST-relaxation dispersion. Filled and open pulses indicate 90° and 180° rf pulses, respectively. The 90° pulses used in the initial INEPT transfer are PC9 [127] pulses, all other selective pulses are BURP pulses [42].

For slow exchange processes (a few 100 ms to a few s) NMR exchange spectroscopy experiments (EXSY) are exploiting the long longitudinal relaxation times T_1 to monitor the exchange process. The principle is similar to NOE spectroscopy. The experiment is correlating the two different states during a mixing time period (figure 29a) and the exchange rates describing the kinetics can be fitted out of the peak intensities and their ratios. Generally ^{15}N spin correlations are used as dipolar mediated NOE effects can be safely neglected due to low γ and large distances. EXSY allows an easy measurement of the chemical shift differences directly from the peak correlation patterns and the populations of both states from the ratios of the peak intensities. Further, it can be used for assignment of the different forms of the protein, if the chemical shifts of one state are already known.

Relaxation dispersion experiments are another option to analyze exchange processes. These experiments are sensitive to a larger timescale (a few 100 µs to a few 100 ms) [100, 122, 176], and cover therefore a broader range of dynamics. An effective transverse relaxation rate $R_{2,eff}$ can be monitored by a repetition of 180° pulses (CPMG train) applied at different frequencies (figure 29b), ν_{cpmg}. The applied pulses partly refocus by the stochastic exchange process dephased transverse magnetization. Relaxation dispersion profiles defined by the dependance of the measured transverse relaxation rate on the frequency of the CPMG train, $R_{2,eff}(\nu_{cpmg})$ can be fitted to

extract the exchange rates, populations and chemical shift differences(in slow exchange). The group of Lewis Kay has been very successful in this context and showed that it can even be used to extract structural constraints of low populated states that are invisible in the spectrum [52]. Relaxation dispersion data also allows to characterize exchange processes involving more than two states, but interpretation and fitting becomes rather difficult, thus additional data is needed.

Both types of experiments are complementary. Especially in the case where two (or more) distinct mechanisms contribute to the exchange process. EXSY determined exchange rates can for example be used to separate slower processes from the faster ones, only visible in the relaxation dispersion data.

One drawback of CPMG type experiments is that the exchange contributing to the transverse relaxation is generally small. Very good signal to noise is needed for clean fitting of the data, the recording time of relaxation dispersion experiments can be very long, and better sensitivity is needed. Therefore we implemented the BEST type pulse sequence optimization into a ^{15}N-^{15}N BEST-EXSY (figure 29a) and a ^{15}N BEST-CPMG (figure 29b) type experiment. This work is still ongoing but preliminary results have already been obtained that will be briefly presented in the remaining of this chapter. To demonstrate that the BEST sequences perform better in terms of sensitivity we have experimentally compared the BEST relaxation dispersion experiment with the standard CPMG experiment [100, 176] using a 2 mM Ubiquitin sample at pH 6 and 25°C. All experiments have been carried out on a 600 MHz Varian spectrometer equipped with a cold probe. Figure 30 shows the sensitivity curves obtained under this condition. Both curves were normalized for equal experimental time. Similar to other BEST experiments, the BEST CPMG experiment performs better at high repetition rates with a maximum reached at a scan time of about 600 ms. The overall gain obtained comparing sensitivity under optimal conditions for each experiment is about 1.5, similar to the results obtained for other BEST type sequences.

At pH 2 and 45°C ubiquitin is showing exchange between an unfolded and folded form. We used this model to setup experimental conditions running relaxation dispersion and EXSY experiments using the BEST versions of the pulse sequences. It has to be mentioned that the major drawback of this type of approach is that the high repetition rate of 180° ^{15}N pulses during the CPMG train results in severe probe heating, thus the compensation pulse at the beginning of the sequence T_{comp}, as well as the CPMG pulse

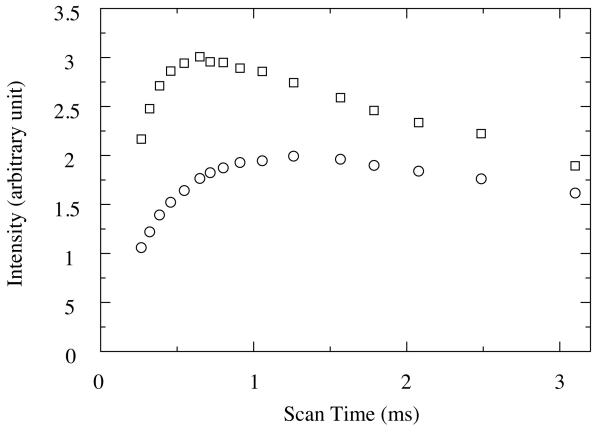


Figure 30 – Average sensitivity of a standard CPMG relaxation dispersion experiments (circles) and the BEST CPMG relaxation dispersion experiment (squares) as a function of scan time. The sensitivity curves were obtained from integrating 2D spectral intensities measured for different recycling delays, normalized for equal experimental time.

train itself, have to be carefully calibrated to obtain clean dispersion profiles. Also care has to be taken to avoid damaging of the cryo-probe. To ensure a good calibration of the experimental setup, the 2 mM ubiquitin sample has been used to verify the quality of the $R_{2,eff}$ profiles. As ubiquitin does not show exchange at pH 6 and 25° the profiles obtained have to be flat and should not depend on the CPMG frequency. Therefore we carefully checked dispersion curves measured at experimental conditions with a recycle delay of 750 ms. After appropriate calibration of the pulse sequence, flat dispersion profiles were obtained. Figure 31a shows two examples of dispersion profiles obtained for the sample showing exchange and for the sample at pH 6 and 45°C (without exchange).

Experiments carried out at 45°C showed nearly equally distributed populations of both states. Figure 32a shows an ^{15}N-^{1}H correlation spectrum obtained at 25°C. A single form of the protein is visible. Figure 32b in contrast shows the same spectrum recorded at a temperature of 45°C. It is clearly visible that a second form of the protein is appearing showing resonance dispersion typical for disordered proteins. The folded form of Ubiquitin at pH 2 has been completely assigned using BATCH [93], a tool developed in the laboratory for automatic backbone assignment based on covariance calculations. The EXSY sequence shown in figure 29a allowed us then to achieve a complete assignment of the unfolded form of the protein

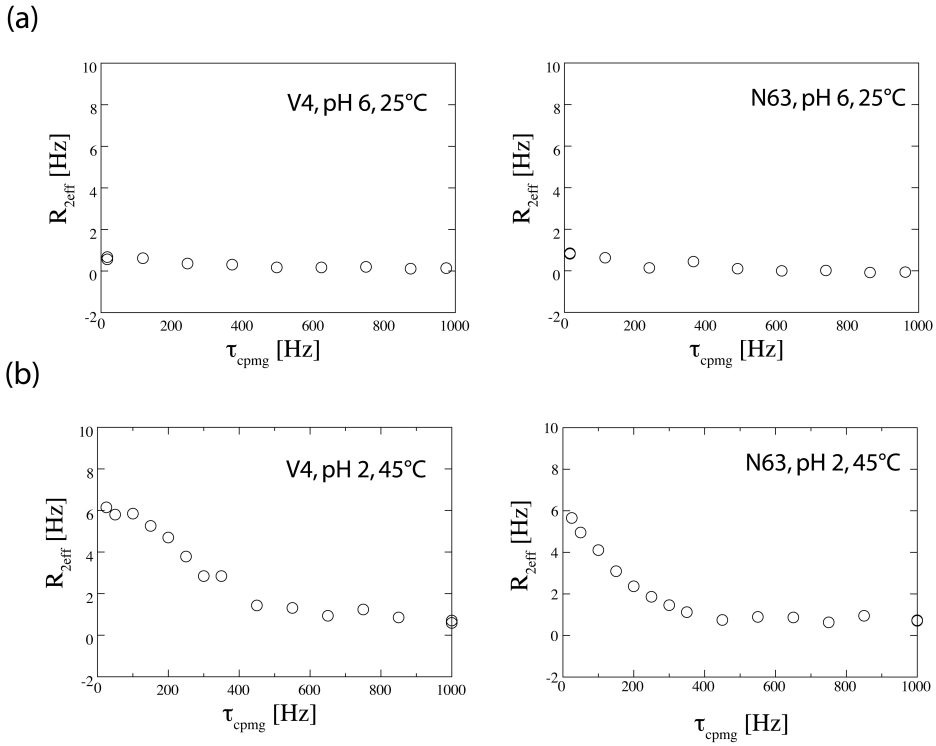


Figure 31 – (a) Flat relaxation dispersion profiles of ubiquitin at pH 2 at temperatures of 25°C. (b) Relaxation dispersion profiles of the same residues recorded at a temperature of 45°C.

following the correlation pattern between both forms in the spectrum (see figure 33). At the same time the chemical shift differences ω_{AB} between both states and their population distribution can be extracted from the spectrum. The high quality of the EXSY spectrum can be appreciated in figure 33 showing the ^{15}N-^{15}N projection of the spectrum. EXSY data has been recorded using a mixing time of 400 ms. A set of relaxation dispersion profiles has been recorded. Some representative examples are shown in figure 31b.

4.4 CONCLUSION

The BEST type pulse sequence optimization is based on the implementation of band selective pulses, to achieve longitudinal relaxation enhancement. To facilitate the pulse sequence design we have introduced simplified schemes for various amplitude-modulated pulse shapes commonly used for band-selective spin manipulation in NMR experiments of macromolecules. These

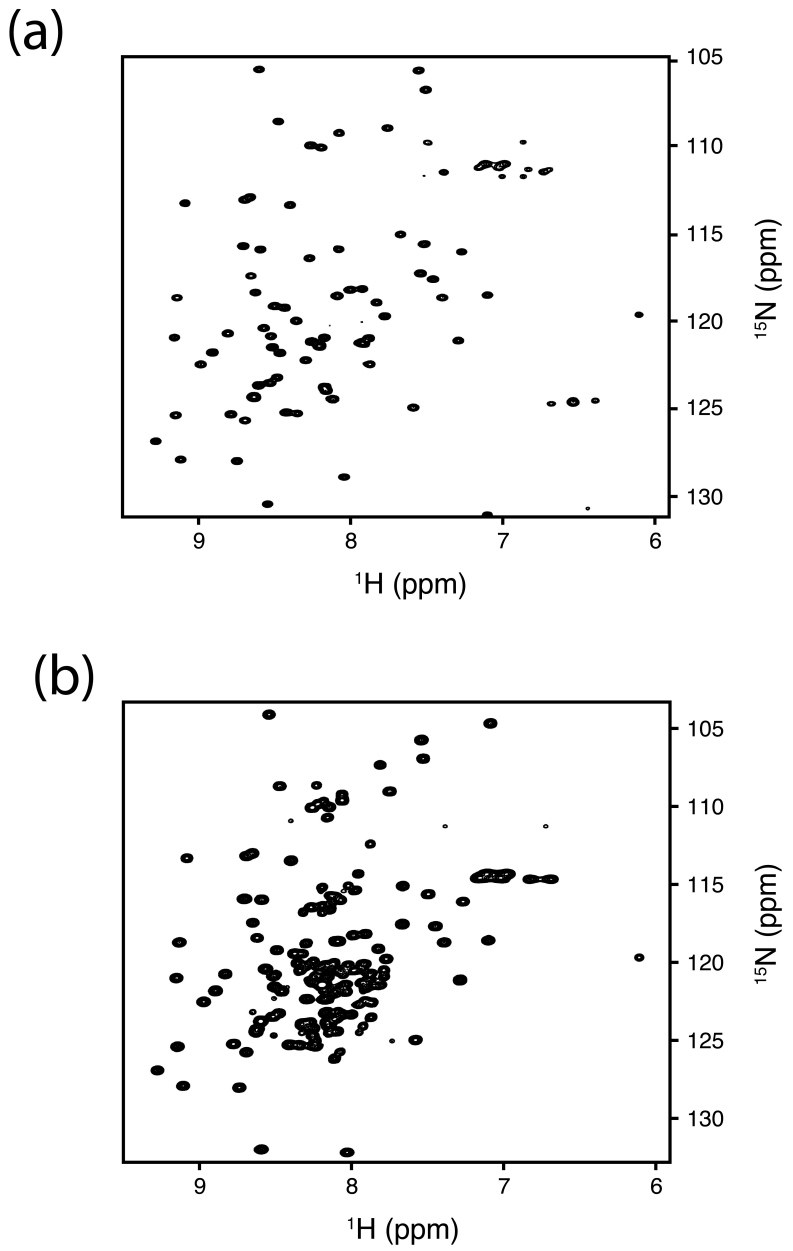


Figure 32 – ^{1}H-^{15}N BEST-HSQC spectra recorded on 0.5 mM ubiquitin at pH 2 at temperatures of (a) 25°C and (b) 45°C, respectively.

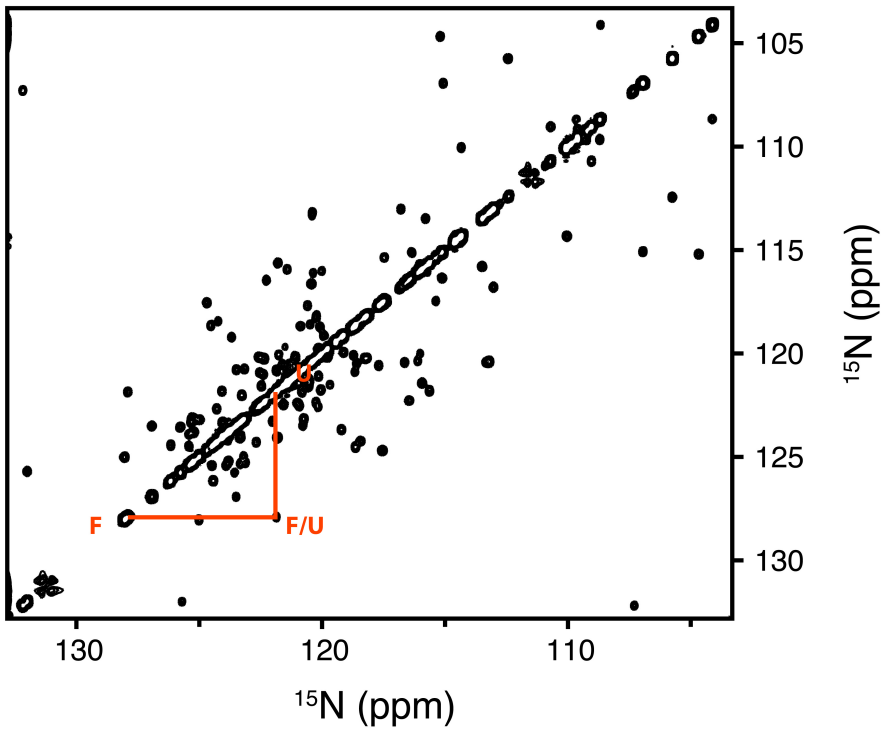


Figure 33 – BEST-EXSY ^{15}N-^{15}N projection of 3D EXSY of a 0.5 mM ubiquitin sample at pH 2 and 45°C. A mixing time of $\tau_m = 400$ ms has been used. The red lines indicate a correlation pattern between the unfolded and the folded form of the protein

binary schemes, consisting in a suite of evolution delays and ideal rf pulses, enable easy use of such pulse shapes in the conception of complex multi-pulse sequences, and ensure their optimal performance, as demonstrated here for longitudinal-relaxation-enhanced BEST-HSQC and BEST-TROSY experiments. Furthermore, we introduced BEST EXSY and BEST CPMG relaxation dispersion experiments allowing the study of kinetic processes in the order of µs to a few 100 of ms. The optimized pulse sequences yield a significantly higher sensitivity that is crucial to this kind of experiments. We also demonstrated for the example of ubiquitin in exchange between an unfolded and folded form that high quality relaxation dispersion profiles can be obtained. A preliminary analysis of our data indicates the presence of faster exchange processes for part of the residues. Further NMR data, e. g. relaxation-dispersion at a second magnetic field strength are required for a more quantitative interpretation of the folding/unfolding processes present in ubiquitin at acid pH and high temperature.

Part III

INTRODUCTION TO THE BACTERIAL CELL WALL

5

THE BACTERIAL CELL WALL

Situated at the barrier between the cell and its environment the bacterial cell wall is essential for bacteria and their survival. It is implied in many filtering and transport mechanism and represents thereby the Achilles heel of bacteria. This necessity of the cell wall to the bacterial survival is exploited by the nearly totality of all antibiotic treatments.

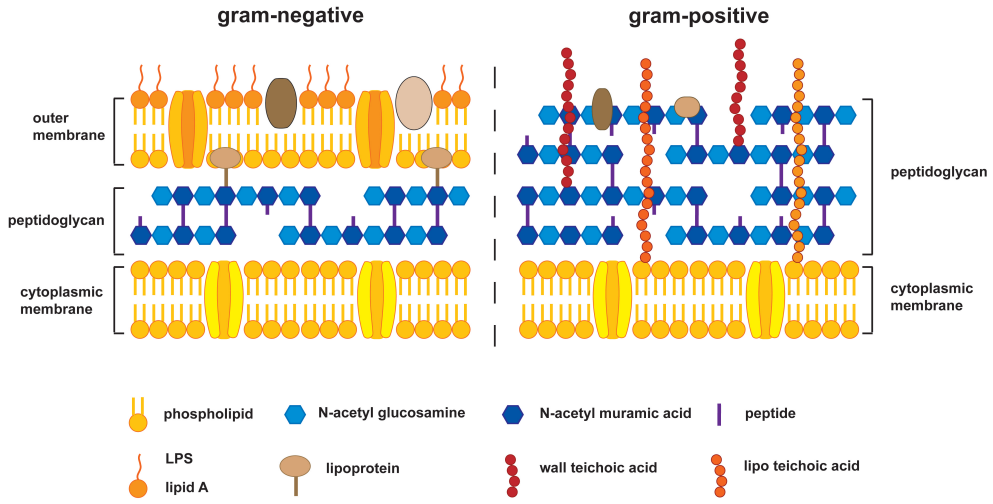


Figure 34 – Schematic representation of the cell wall components of gram-negative and gram-positive bacteria. The structural illustration of the peptidoglycan is purely schematic and doesn't give any information about its structural arrangement

Roughly we distinguish between two types of bacteria, gram-negative bacteria with *Escherichia coli* as a very good established model bacteria, and gram-positive bacteria with *Bacillus subtilis* being one of the most studied examples. Other examples of Gram-negative bacteria are proteobacteria, cynobacteria, spirochaetes, green sulfur and green non-sulfur bacteria. On

the side of Gram-positive bacteria well known genera are *Bacillus, Listeria, Staphylococcus, Streptococcus, Enterococcus, Diplococcus, pneumoniae, and Clostridium*.

The most obvious difference between both types of bacteria is the missing outer membrane in Gram-positive bacteria, where a network of anionic polymers, the teichoic acids, replace their function, and the drastic difference in thickness of the peptidoglycan, where it is much thinner in Gram-negative bacteria than in Gram-positive bacteria.

The classification is historical and based on the negative or positive response of the cell wall to gram staining, a coloration employed since long time where the response is mainly due to the thickness of the peptidoglycan layer of Gram-positive bacteria and their lack of a outer membrane allowing staining (1884 by Christian Gram [48, 164]). A shematic representation of the differences in the cell wall of both classes is shown in Figure 34. It has to be mentioned that there are also few bacteria that are neither qualified as Gram-negative nor as Gram-positive such as the very pathogen *Mycobacterium tuberculosis* because of chemical differences to both classes. Generally both classes contain pathogen species. In Gram-negative bacteria the presence of a capsule is increasing the virulance and lipopolysaccharides attached to the outer membrane of Gram-negative bacteria can cause severe inflammation up to a septic shock. The peptidoglycan of Gram-positive bacteria is more vulnerable to lysozyme digestion due to its open exposure also resulting in a higher susceptibility of Gram-positive bacteria to antibiotics.

The work presented in this part of the thesis is focused on the chemical characterization of the cell wall and its interaction with proteins and metal ions. In this context different parts of the cell wall and some of their function will be discussed in this chapter. The focus will be on the peptidoglycan and the teichoic acids in Gram-positive bacteria as they are the center of interest in this study.

5.1 PEPTIDOGLYCAN

The peptidoglycan (also called murein) is the major component of the cell envelope of most bacteria. In Gram-negative bacteria it makes up around 10% of the dry weight while it can go up to 90% for Gram-positive bacteria. It protects the cell against environmental stress like the turgor pressure [121] and is maintaining cell's shape [174, 178]. Unique to the bacterial world

it is the key target for most antibiotic treatments, more precisely for the two major antibiotic classes, the β-lactams and glucopeptides, inhibiting its biosynthesis [79]. Fragments released from the peptidoglycan are also recognized by the immune system provoking a response of the innate immune system [15, 49, 114, 131]. While chemical composition and structure of fragments have been established, and several enzymes involved in peptidoglycan assembly have been isolated, the 3-dimensional organization of this biopolymer is still debated [175]. Orientations of the glycan strands parallel [78, 173] or perpendicular [24, 118] to the cell membrane have been proposed (see section 3-dimensional structure of the peptidoglycan). Its biosynthesis and growth mechanism remain an area of active research.

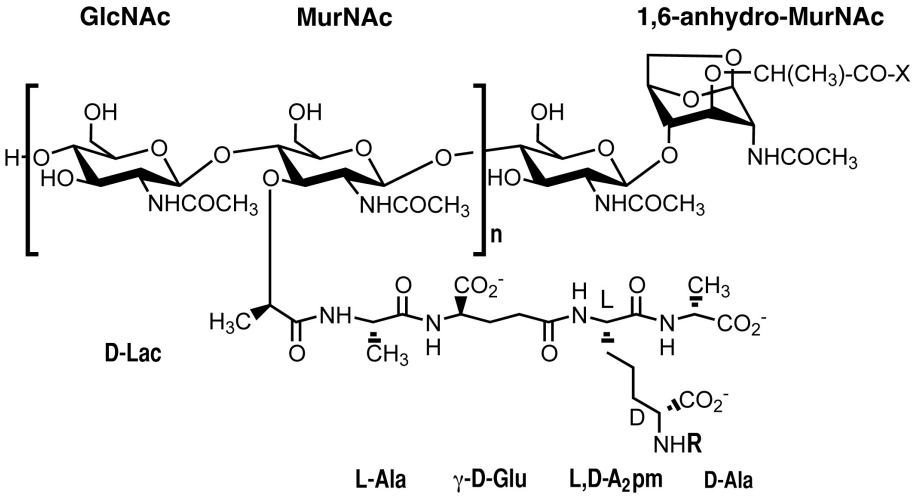


Figure 35 – Repeating unit of the peptidoglycan of *E. coli*. X indicates a peptide strand linked to the lactyl group of 1,6-anhydro-MurNAc, R stands for the cross-linked peptide attached to a neighboring glycan strand of peptidoglycan.

The primary structure of peptidoglycan consists of glycan strands of alternating β(1-4) linked N-acetylmuramic acid (MurNAc) and N-acetylglucosamine (GlcNAc) residues cross-linked by short peptides which are attached to the lactyl group of the N-acetylmuramic acid (figure 35). They are terminated with a 1,6-anhydro-MurNAc instead of a MurNAc residue [47]. A lactyl group is attached to the C_6 of the MurNAc carying a pentapeptide cross-linking the glycan strands, thus building up a mesh like structure surrounding the cell. While the chemical composition of the sugars does not vary the peptide composition and the chemistry of the cross-linking can be significantly different between bacterial strains and species [174]. The glycan strand length varies depending on the species and growth conditions. As an

example, the length of the glycan strands in the Gram-negative model bacterium *Escherichia coli* vary from 1 to 80 disaccharide units with an average of approximately 21~35 units [47, 54]. In contrast, the Gram-positive *Bacillus subtilis* has glycan strands of up to 5000 disaccharide units [56]. There are some general differences between peptidoglycans of the two different Gram classes which will be briefly discussed in the following:

Gram-negative bacteria

The peptidoglycan of Gram-negative bacteria is situated in the periplasmic space between the cytoplasmic membrane (or inner membrane) and the outer membrane. Braun's lipoprotein is bond covalently to the peptidoglycan linking it to the outer membrane. As mentioned a major difference between both classes is the difference in thickness of the peptidoglycan. To give two examples the thickness of the peptidoglycan of *E. Coli* was detected to be 6.35 nm and the one of *P. aeriguniso* is 2.41 nm [109]. Both values where measured by cryo-transmission electron microscopy. These values also illustrate the still ongoing debate if there is a layered structure, and if Gram-negative bacteria contain one of them or several. With a thickness nearly three times higher as the peptidoglycan of *P. aeriguniso* the peptidoglycan of *E. Coli* seems to be thicker than it should be for a completely mono-layered organization (see section 3-dimensional structure of the peptidoglycan).

Gram-positive bacteria

The thick peptidoglycan is anchored to the inner membrane. Similar to the periplasmic space in Gram-negative bacteria a so called inner wall zone (IWZ) has been revealed by cryo-electron microscopy [111]. It is placed in between the peptidoglycan with its teichoic acids (outer wall zone OWZ) and the inner membrane and has similar composition and function than the periplasm [110]. The peptidoglycan of Gram-positive bacteria is also generally much thicker than its Gram-negative counterpart. Two examples are the teichoic acid containing peptidoglycan of *B. subtilis* and *S. aureus* with 19 nm and 28 nm, respectively [110, 111]. This values were measured in the intact cell under the influence of the turgor pressure, isolated peptidoglycan can be thicker showing that the high pressure is compressing it. Another interesting point is that the peptidoglycan-teichoic acids containing cell wall of *B. subtilis* has the same thickness as the isolated peptidoglycan while the peptidoglycan of *S. aureus* is reduced in thickness after removal of the teichoic acids. This indicates a significant difference in the wall teichoic acids/peptidoglycan organization between both species.

5.1.1 *Chemistry of the peptidoglycan's peptide stems and mechanisms of cross-linking*

While the composition of the glycan strands is mainly varying in length but not significantly in their composition, the chemistry of the peptide and the cross-linking between neighboring glycan strands can differ significantly between bacterial species.

Chemistry of the peptide stems

In most gram-negative bacteria the chemical composition of the pentapeptide attached to the MurNAc over the D-lactoyl group is L-Ala-γ-iso-Glu-L,D-mesoA_2pm-D-Ala-D-Ala [47]. In Gram-positive bacteria the mesoA_2pm is in the majority of species replaced by a L-Lys group. The D-Alanine in position 5 is lost after cross-linking. Nevertheless the chemical composition can vary among different species and strains [173, 174] (see table 1). The elementary repeating units of the peptidoglycan in *E. coli*, *B. subtilis* and *S. aureus* are shown in figure 36. These variations can be separated in two classes [174]. Firstly, variations occurring during the assembling of the peptidoglycan, due to the specificity of the Mur ligases, an enzyme family responsible for the biosynthesis of the peptidoglycan precursor. Secondly, modifications posterior to synthesis often due to growth conditions e.g. in *B. subtilis* growth dependent amidation of meso-A_2pm occurs after the action of Mur ligases [3, 177].

Figure 36 – Schematic representation of different peptide cross-links found in the peptidoglycan networks of (b) *E. coli*, *B. subtilis* and (c) *S. aureus*

Going through the different positions of the peptide stem the MurC ligase is attaching the first amino acid to the UDP-GlcNAc-MurNAc precursor. This position is in almost the totality of bacteria a L-alanine, only in very rare cases it can be replaced by a glycine or L-serine residue. The same

is true for position two where the MurD ligase is adding in all bacteria D-isoglutamate acid which can be modified posterior to D-isoglutamine or threo-3-hydroxuglutamate.

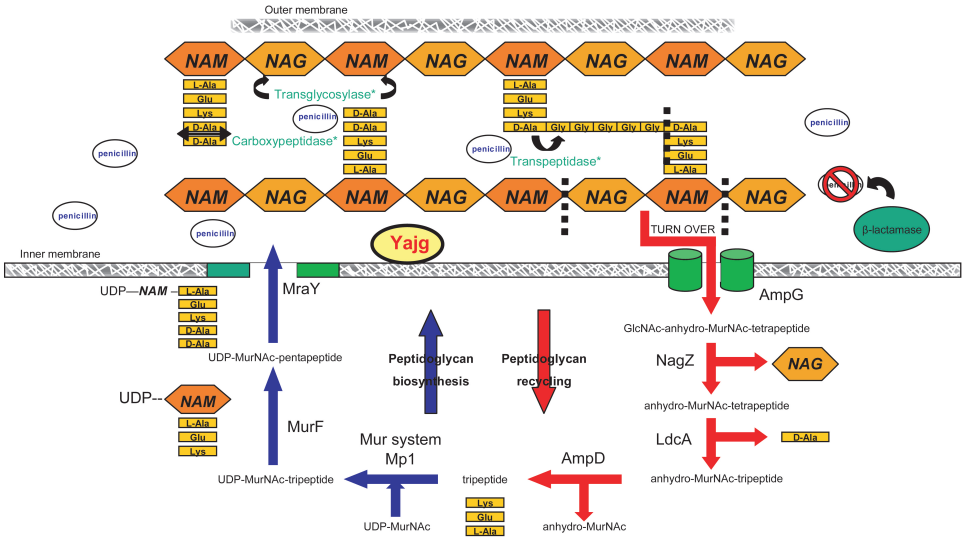


Figure 37 – Schematic representation of the peptidoglycan biosynthesis

The most significant variations occur in position three. Common to all of them is that the MurE ligase is adding a diamino acid for later cross-linking, generally meso-A_2pm for Gram-negative bacteria and L-lysine for Gram-positive. Nevertheless following modifications or amino acids can also be found in position three of different species: Amidated meso-A_2pm, LL-A_2pm, L-lysine/D-lysine, L-lysine/L-ornithine, L-ornithine, meso-lanthionine, L-2,4-diaminobutyrate, L-hormoserine, L-alanine, L-isoglutamate, 2,6-diamino-3-hydroxypimelate, L-5-hydroxylysine and N^{γ}-acetyl-L-2,4- diaminobutyrate.

Finally the D-alanine on position four and five is added as a dipeptide by MurF. While position four is not known to vary, position five can be replaced by a D-serine or D-lactate in the context of resistance against the antibiotic vancomycin.

Position	Residue encountered	Examples
1	L-Ala	Most species
	Gly	Mycobacterium leprae, Brevibacterium imperiale
	L-Ser	Butyribacterium rettgeri
2	D-Isoglutamate	Most Gram-negative species
	D-Isoglutamine [1]	Most Gram-positive species, Mycobacteria
	threo-3-Hydroxyglutamate [1]	*Microbacterium lacticum*
3	meso-A2pm	Most Gram-negative species, Bacilli, Mycobacteria
	L-Lys	Most Gram-positive species
	L-Orn	*Spirochetes, Thermus thermophilus*
	L-Lys/L-Orn	*BiÞdobacterium globosum*
	L-Lys/D-Lys	*Thermotoga maritima*
	LL-A2pm	*Streptomyces albus, Propionibacterium petersonii*
	meso-Lanthionine	*Fusobacterium nucleatum*
	L-2,4-Diaminobutyrate	*Corynebacterium aquaticum*
	L-Homoserine	*Corynebacterium poinsettiae*
	L-Ala	*Erysipelothrix rhusiopathiae*
	L-Glu	*Arthrobacter J. 39*
	Amidated meso-A2pm [1]	*Bacillus subtilis*
	2,6-Diamino-3-hydroxypimelate	*Ampuraliella regularis*
	L-5-Hydroxylysinew	*Streptococcus pyogenes*
	N^{γ}-Acetyl-L-2,4-diaminobutyrate [1]	*Corynebacterium insidiosum*
4	D-Ala	All bacteria
5	D-Ala	Most bacteria
	D-Ser	*Enterococcus gallinarum*
	D-Lac	*Lactobacillus casei, Enterococci with acquired resistance to vancomycin*

Table 1 – Amino acid variations in the peptide stem taken from "Peptidoglycan Structure and Architecture" by Vollmer et al [174]

Mechanism and variation of cross-linking

Peptide cross-linking most often occurs between the amino acid at position 3 [meso-A_2pm for *E. coli* and *B. subtilis*, L-Lys(Gly5) for *S. aureus*] of one peptide strand and the D-Ala at position 4 of another peptide strand with release of the D-alanine in position 5 (3-4 cross-linking). This cross-linking is catalyzed by the D,D-transpeptidases also called Penicillin Binding Proteins (PBP) because of their affinity to β-lactam antibiotics like G-penicillin. This enzymes are specific to the D,D bond of the terminal D-Ala-D-Ala peptide. In the transpeptidation reaction the alanine in position four is covalently attached to the diamino acid in position three of the neighbouring strand and the D-alanine in position five is cleaved off. Figure 38 is showing transpeptidation on the example of a Gram-negative bacteria like *E. coli* (or the Gram-positive bacteria *B. subtilis*).

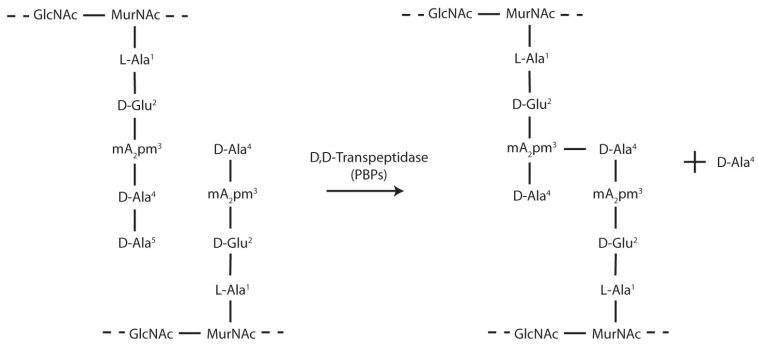


Figure 38 – Schematic representation cross-linking for the example of a D,D-Transpeptidase inducing cross-linking between the amino acid in position 3 of the donor strand and 4 of the acceptor strand with cleavage of the D,D bond between the D-Alanine in position 3 and 4 of the acceptor strand.

In a weaker percentage transpeptidation can also occur between the diamino acid in position three of one strand to the diamino acid of the neighbouring strand (3-3 cross-linking). This transpeptidation is also a mechanism of resistance against antibiotic stress. It is done by L,D-transpeptidases [102, 105] which are not sensible to most of the antibiotics because they are specific to the L,D-bond between the L-diamino acid and the D-alanine in position three and four, respectively. An example of L,D transpeptidation in Gram-negative bacteria is shown in figure 39. Nevertheless 3-3 cross-linking percentage is rather low but raises in the stationary phase or in β-lactam resistant strains. The cross-linking between the amino acids in both peptide strands involved can be direct or over a peptide bridge which can vary between one and seven amino acids [174]. Examples for peptide bridges

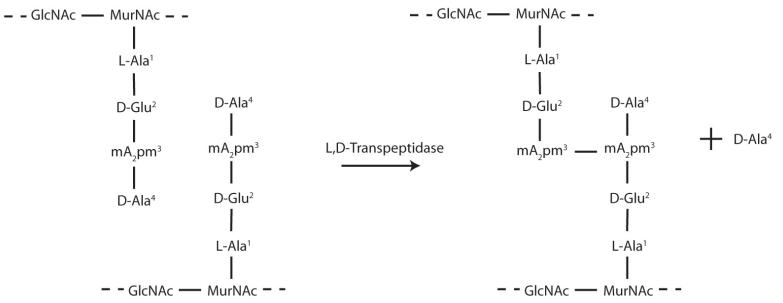


Figure 39 – Schematic representation of cross-linking for the example of a L,D-Transpeptidase inducing cross-linking between the amino acid in position 3 of the donor strand and 3 of the acceptor strand with cleavage of the L,D bond between the L-Lysine in position 3 and the D-Alanine in position 4 of the acceptor strand.

involving one peptide is *E. faecium* where the cross-link is done involving one asparagine or aspartate residue (figure 40a) while in the peptidoglycan of *S. aureus* the cross-linking is done over a glycine-pentapeptide bridge (figure 40b). A more exotic example is the cross-linking in *M. luteus* where a large part of one peptide stem is detached from the MurNAc residue by an amidase reaction and then involved in forming a bridge between residue in position three and four(figure 40c). The degree of cross-linking is species dependent as well as growth-phase (exponential versus stationary phase) dependent. In the peptidoglycans from *E. coli* and *B. subtilis* approximately 40-60% of the peptides are part of cross-links, whereas over 90% of the peptides are cross-linked in *S. aureus* [17].

It should be mentioned that there exist another form of cross-linking where the cross-linking is done between the D-isoglutamate in position two and the D-alanine in position four. This type of cross-linking is rare and appears only in coryneform bacteria.

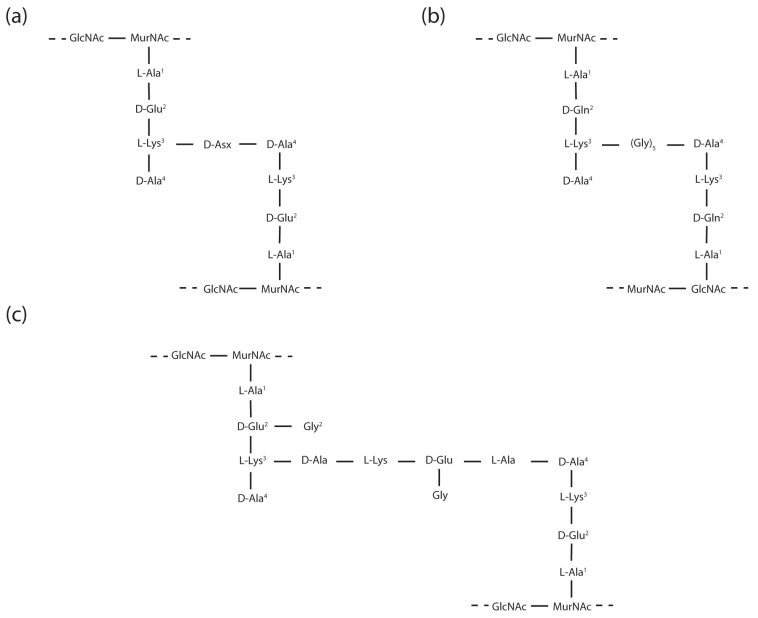


Figure 40 – (a) 3-4 cross-link over a penta-Glycine bridge in *S. aureus*. (b) 3-4 cross-link involving a peptide stem bridge on the example of *M. luteus*. (c) 3-4 cross-link involving a D-Asp or D-Asn residue as bridge on the example of *E. faecium*

5.1.2 *3-dimensional structure of the peptidoglycan*

The 3-dimensional structure of the peptidoglycan is highly debated [175]. Contrary to similar polymers like chitin or cellulose, it has not been possible for now to solve an atomic resolution structure of the polymer. This lack of information can be explained by the size and non crystallin structure of the peptidoglycan. Nevertheless, several models have been proposed based on physical properties and the chemical composition of the peptidoglycan. Principally they can be divided into two types of models. The layered model, with glycan chains running parallel to the membrane and the scaffold model with glycan chains perpendicular to the membrane.

Until now the most accepted model has been the layered model where the glycan strains are parallel to the cell surface (figure 42a) and has been recently challenged by the so called scaffold model where the glycan chains are perpendicular to the surface (figure 42b).

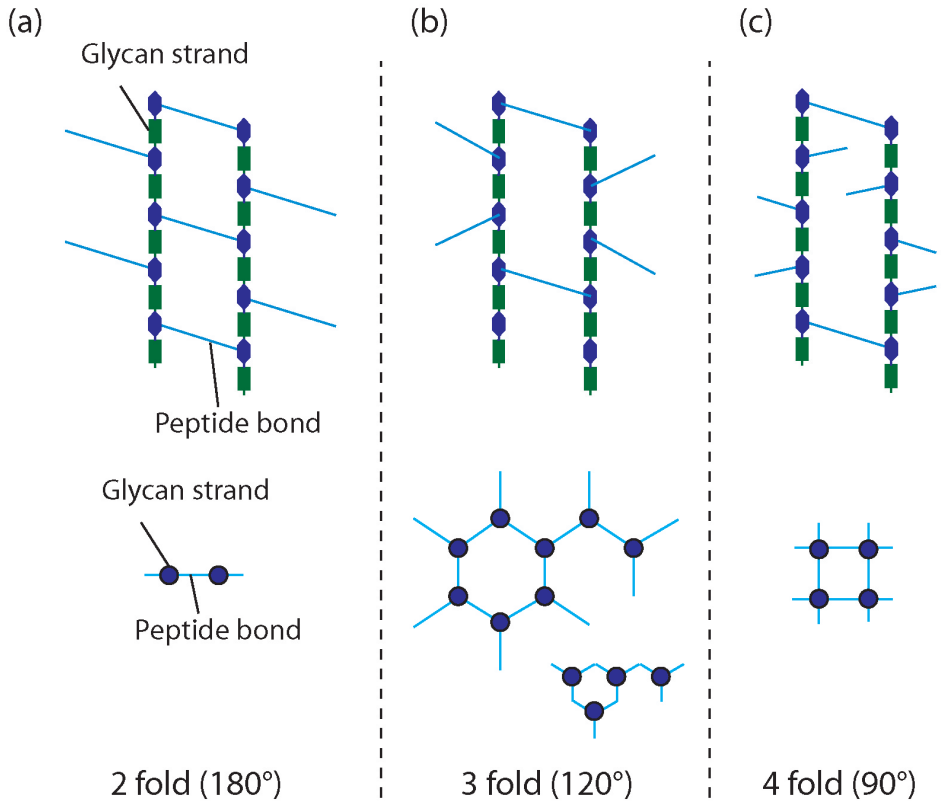


Figure 41 – Schematic representation of different fold types in the peptidoglycan, depending on the torsion angle (in brackets) between two succedent peptide stems along the glycan chain. (a) 2-fold structure like (chitin), (b) 3-fold (peptidoglycan, scaffold model), and (c) 4-fold layer (horizontal layer model).

Layered model

Historically the layered model has been the first one proposed [69, 70, 118]. It is mainly based on structural predictions and its similarity to chitin in the primary structure of the glycan strains. While in chitin two following GlcNAc residues are turned by 180° respectively to each other (figure 41a), it has been shown that the MurNAc residue in the Peptidoglycan only allows less rotation (figure 41b and c). The actual consideration is that the pentapeptides are arranged in a helical pattern where four disaccharides are needed for one complete turn (figure 41c) [87]. This means that every second peptide is laying in the same plane and is therefore able to form a cross-link to neighboring glycan strands. Therefore a mono-layered peptidoglycan could only be cross-linked at about 50% because the other half of pentapeptides would point up and down of the plane. This is in a good agreement with

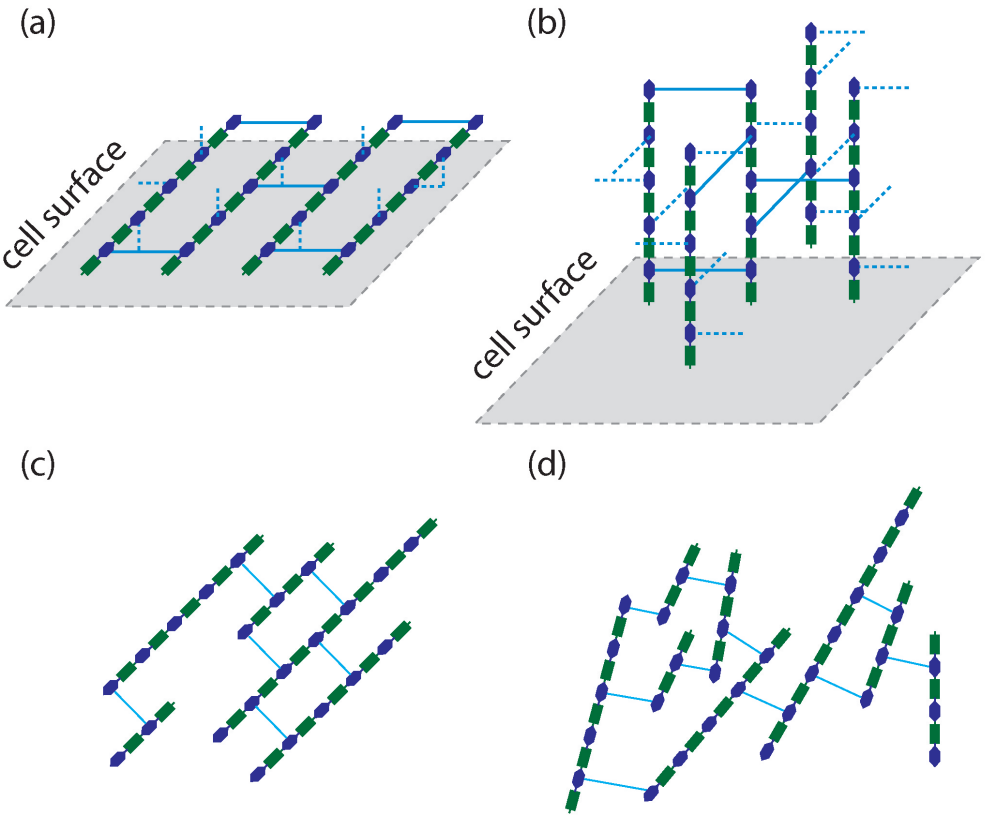


Figure 42 – (a) The layered model with glycan strands going parallel to the cell surface. (b) Scaffold model with glycan strands going vertical to cell surface. (c) Schematic representation of the glycan strands being homogeneously organized in regard to each other. (d) Schematic representation of the glycan strands being randomly orientated in regard to each other.

the cross-linking pourcentage measured on the mono-layered peptidoglycan of *E. coli* of about 40% to 50% [46, 47].

Another point is the orientation of the glycan chains. They run (i) parrallel and along the long axis of the cell, (ii) parrallel and perpendicular to the cell (42c for i and ii), (iii) parallel and along helices around the cell, (iv) in areas of parallel strands together with areas of random orientation and (v) randomly without any order (42d).

Scaffold model

The scaffold model is more recent and has been mainly introduced by Dmitriev et al. [22–24] and recently also by the group of Mobashery [118]. In the scaffold model glycan strains are running in a plane perpendicular to the

cell surface and are cross-linked by pentapeptides laying in a plane parallel to the membrane. The model has been based on published data on the cell wall assembly and computations. In the case of the group of Mobashery, NMR distance and torsion angle constraints have been measured on a small fragment of the peptidoglcyan (two succedent dissacharides). They conclude that the torsion angles in the helical pattern of the glycan strands are close to 120° giving a 3-fold organization (figure 41b). Nevertheless the model is highly debated [175] and contradictory to several experimental findings, especially the well established glycans strain length [47, 54] being many times longer then the peptidoglycan thickness.

5.1.3 *Antibiotics as an inhibitor of the peptidoglycan biosynthesis*

With the discovery of penicillin by Alexander Fleming in 1928 β-lactam antibiotics became the most powerful drug against bacterial infections. Initially mostly efficient against Gram-positive bacteria modern penicillin derived antibiotics are nowadays as efficient against Gram-negative bacteria and are indispensable in clinical treaments.

Their mode of action is the inhibition of cross-linking by the D,D-transpeptidases (Penicillin Binding Proteins) and therefore of the peptidoglycan biosynthesis. Mimicking the D-alanine-D-alanine peptide (figure 43) terminus of the pentapeptide of the nascent peptidoglycan β-lactams have a high affinity for PBPs and bind covalently to the active site serine of the enzyme, thus deactivating it. This prevents the cross-linking in the peptidoglycan layer leading to a degradation of the integrity of the cell wall. As additionally peptidoglycan precursors trigger the activation of cell wall hydrolases, needed for the incorporation of the precursors in the peptidoglycan layer, the increase in precursor concentration due to the inhibition is resulting in an even increased bactericide.

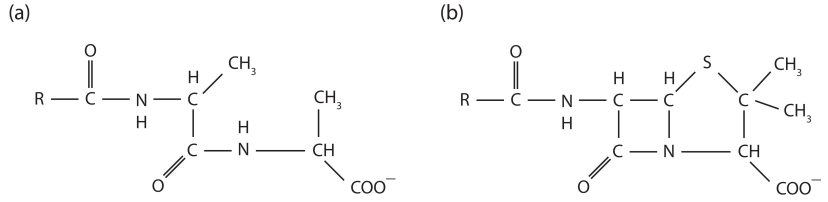


Figure 43 – Comparison between (a) the D-Ala^4-D-Ala^5 peptide and (b) the penicillin core structure

As various bacterial species are developing strong resistances against antibiotic treatments the research for new antimicrobial agents is becoming urgent. Therefore a better understanding of the mechanisms of resistance and the cell wall biosynthesis in general is highly needed. Up to date three major mode of resistance are targeted. The first is an enzymatic hydrolysis of the β-lactam ring making the antibiotic useless, the second a mechanism involving active pumps rejecting the antibiotic, and finally altered PBPs having a much lower affinity for β-lactams. Mainly the mechanism of the latter resistance is not well understood yet and is under heavy research. In the last years L,D transpeptidase got more and more in the focus of research [103]. As they are not acting on the D,D bond of the donor they are expected to lack inhibition by β-lactam antibiotics which are analogues of the D-Ala^4-D-Ala^4 peptide terminus. Therefore they can act as an escape mechanism upon antibiotic stress or in resistant bacteria forming 3-3 cross-links to maintain peptidoglycan integrity [104].

5.1.4 *The role of the peptidoglycan in the innate immune system*

The innate immune system is one of the first responses from the host to the presence of a pathogen. It is common to all plants and animals and is dominant in plants, fungi and insects as in primitive multicellular organisms. Contrary to the adaptive immune system it does not provide long-lasting adaptive protection to the host but acts in a more direct way and provides immediate defense against infection. It triggers a inflammatory response producing chemical factors which attract phagocytes which can then trigger other parts of the immune system.

The innate immune system is activated by the recognition of pathogen associated molecular pattern (PAMPs). These PAMPs are molecules "foreign" to the host organism. The peptidoglycan is one of the major components recognized by the immune system in case of bacterial infection and several receptors have been found to be involved in the immune response to pathogen bacteria. Proteins involved in the recognition of such specific molecules are called pattern recognition molecules (PRMs). Well known PRMs involved in peptidoglycan recognition are the Toll like receptor 2, Nod 1, and Nod2 [43]. In mammalians, Peptidoglycan Recognition Proteins (PGRPs) [49] are also sensitive to the peptidoglycan and act directly as bactericide by binding and, in certain cases, hydrolyzing the peptidoglycan. Thus, the mammalian innate immune system exploits the uniqueness and variable composition of peptidoglycan for the recognition of specific bac-

terial pathogens. Other cell wall components recognized by the immune system are lipopolysaccharides (LPS) in Gram-negative bacteria, lipoteichoic acids in Gram-negative bacteria.

5.2 TEICHOIC ACIDS

Teichoic acids (TAs) are present in the cell envelope of Gram-positive bacteria replacing the outer membrane of Gram-negative bacteria. Wall teichoic acids (WTAs) are covalently bound to the N-acetylmuramic acid of the peptidoglycan by a phosphate group, whereas lipoteichoic acids (LTAs) are anchored by a glycolipid to the cytoplasmic membrane. In Gram-positive bacteria lacking capsule and S-layer, like *B. subtilis*, peptidoglycan and TAs are at the interface between the cell and its environment. The polyionic network formed by the TAs has many diverse roles. Of these, the best known relates to the regulation of metal cation homeostasis and the ion trafficking into and out of the cytoplasm [124]. TAs are also implicated in cell adhesion mechanisms through their interaction with mammalian receptors [131]. Weidenmaier et al. demonstrated that this surface-exposed polyanionic polymer plays a key role in nasal colonization by *S. aureus* [179]. It has also been shown that TAs act in the modulation of interactions between the cell surface and endogenous cationic antimicrobial peptides (CAMPs) [136], as well as in the formation of biofilms [180].

WTAs and LTAs are composed of a linkage unit followed by linear poly-(ribitolphosphate) or poly-(glycerolphosphate) chains (44). Several substituents are frequently present on the repeating units where they modulate the TA activities. The best known of these substituents are protonated D-alanyl ester residues that are abundant on the TAs of low- GC bacteria like *B. subtilis*. Their positive charges act to reduce the overall negative charge of the cell wall and thus reduce bacteria sensitivity to CAMP [136]. *S. pneumoniae* which is lacking D-alanyl ester subsituents has phosphorylcholine groups in their place to counter the anionic charges.

5.2.1 *The role of teichoic acids in metal binding in the cell wall*

It has been proposed since the early 70s that teichoic acids are involved in the cation metal binding to the cell wall [5, 61]. Thus, it has been shown by Murray and his coworkers [8] on the cell wall of *B . subtilis* Marburg 168 that its peptidoglycan/teichoic acid network is able to uptake large amounts of

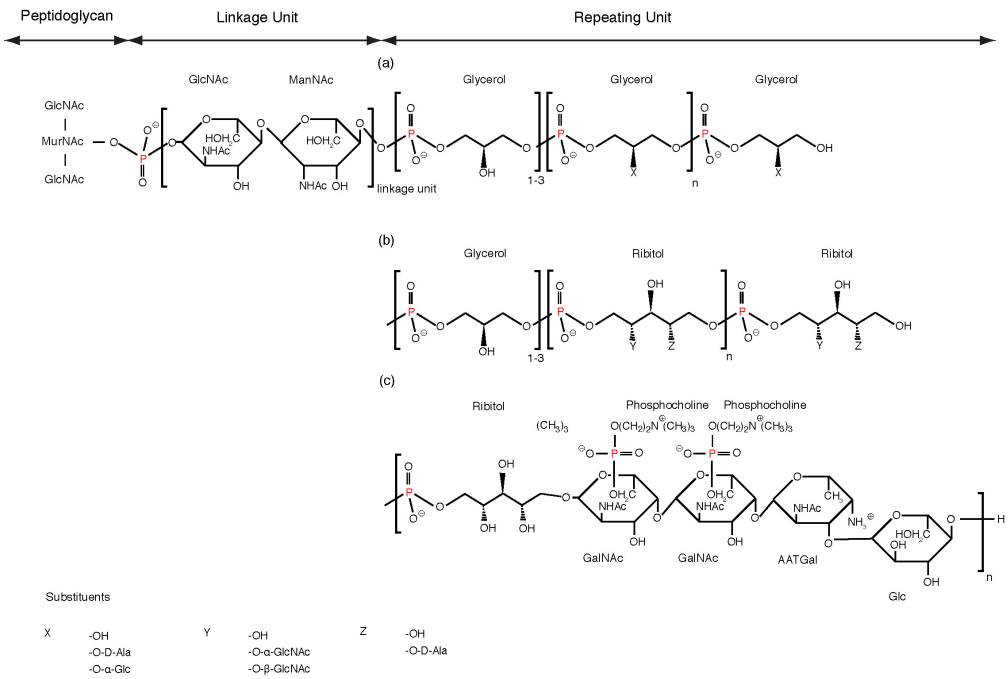


Figure 44 – Chemical structure of WTAs in (a) *B. subtilis*, (b) *S. aureus* and (c) *S. pneumoniae* R6. WTAs are covalently bound at position 6 of the N-acetylmuramic acid of the glycan strand. Possible substituents X, Y and Z are detailed at the bottom of the figure

cations (Na, K, Mg, Ca, Mn, FeIII, Ni, Cu and AuIII) with a strong binding affinity. In this study it has also been shown that bound magnesium ions present the major contribution. They also proposed precociously that the glutamic carboxyl group and the teichoic phosphodiester groups are the major players in binding.

Even if the binding is know since more then 30 years for now, it is still uncertain how the cations are exactly binding to the teichoic acids or the peptidoglycan. Two main binding modes have been proposed. A monodentate binding to one phosphate group while the neighboring phosphoate group creates a contact ion pair with the alanine substituent and a monodentate bridging of Mg^{2+} between two phosphates without a contact ion pair creation. The absence of better structural information on binding is mainly due to the few information available on the exact structure of the cell wall because of its amorphous character and size. While some studies have been done by liquid state NMR on isolated and short teichoic acid chains [139] solid state NMR seems to be an adapted method to study metal

chelation directly on the entire cell wall. This has already been shown by the group of Rice proposing different structural conformation using ^{31}P CPMAS NMR combined with theoretical calculations [181]. Comparing simulated NMR spectra based on energy-minimized structures of 6 possible coordination schemes they propose a anhydrous or hydrated bidentate binding but excluded the monodentate models. While they have shown that solid state NMR can contribute to the topic of metal binding in the cell wall it is still on lyophilized samples of isolated teichoic acids and the role of the peptidoglycan in the binding mechanisms remains uncertain.

5.2.2 *Bacterial adhesion via teichoic acids*

Teichoic acids are thought to be involved in cell adhesion to surfaces where they have been proposed as initial adhesion molecule. Cell adhesion is the first step in biofilm creation which is a structured accumulation of cells forming a protected biological environment within a self-developed polymeric matrix. This biofilm creation is proposed as a four step process where cells in their free floating form (planktonic) first adhere to a living or inert surfaces via the weak van der Waals forces and anchor then in a stronger way by the use of cell adhesion mechanisms. They then form a extracellular polymeric matrix for protection and develop a closely packed architecture. The biofilm can then spread by bacterial release. Biofilms are involved in several chronic infections and are a major problem because of their adhesion to materials like implements, catheters, contact lenses and other medical material.

Information about the mechanisms of adhesion are very valuable for the understanding of biofilm growth and for finding treatments for biofilm elimination. An example of the use of solid state NMR in this field is the work of Rice and Wickham [182] where they were studying lipoteichoic acid adhesion on cellulose, peptidoglycan and TiO_2. Analyzing ^{31}P CP MAS for changes in the CSA and asymmetry parameters and combining it with spin-lattice relaxation ^{31}P $T_{1\rho}$ data they were able to gain further insight in the chemical environment and dynamics of the lipoteichoic acids upon cell adhesion.

Part IV

INTRODUCTION TO SOLID STATE NMR

6

SOLID STATE NMR - HOW TO GET INFORMATION ON THE BACTERIAL CELL WALL

Liquid state NMR has become a very powerful and widely applied tool in structural biology because of its sensitivity to small variations in the local chemical environment of biomolecules and is today without fail one of the main techniques in structural biology. However, it is still limited to medium sized structures of proteins and other biopolymers. Even if there has been striking progress in the domain of NMR of large scale proteins (>30 kDa) [59, 163] and big protein/protein or protein/RNA complexes [38], the information we can get out of proteins and other biopolymers above 80 kDa is still limited, compared to what is feasible on smaller bio-molecular structures. Theoretically, Solid-state NMR (ssNMR) does not have this size limitation. While the often broad, featureless and overlapping lineshapes in ssNMR were a major drawback in the early years, the recent development of fast spinning MAS probes and optimized pulse sequences allows the acquisition of highly resolved spectra. Combinations of spin-decoupling, cross-polarization and magic angle spinning exhibit narrow lines, thus making solid state NMR a very attractive tool for structural biology. As the work presented in this part of the thesis is treating the bacterial cell wall which is one of the biggest biopolymer found in prokaryotic cells, solid-state NMR is the method of choice. The following chapter will introduce the necessary basics and spectroscopic approaches used to study the bacterial cell wall. For further reading on the basics of NMR the reader is referred to [96, 115] and for a detailed description of solid state NMR on polymers to [152].

The physics behind solid-state NMR is the same as for NMR in solution. However, the absence of brownian motion does not average the intrinsically anisotropic contributions as it is the case in solution. While they carry valuable exploitable information about structure and dynamics, they also lead to broadening of the resonance peaks and to a higher spectral complexity (overlapping).

The three anisotropic interactions that largely contribute to the resonance broadening are the dipolar interaction, the chemical shift anisotropy (CSA) and the quadrupolar interaction (for spins higher than $\frac{1}{2}$).

6.1.1 *Dipolar Coupling*

The dipolar interaction can be written in the case of high magnetic fields as,

$$\mathcal{H}_D = d_{jk}(3\cos^2\theta - 1)(3I_{Jz}I_{kz} - \mathbf{I_j}\mathbf{I_k}) \tag{6.1}$$

for the homonuclear case, and as,

$$\mathcal{H}_D = d_{jk}(3\cos^2\theta - 1)2I_{jz}I_{kz} \tag{6.2}$$

for the heteronuclear case, where d_{jk} is the dipole-dipole coupling constant in Hz (all other constants in SI units as well).

$$d_{jk} = -\frac{\mu_0}{4\pi}\frac{\gamma_j\gamma_k\hbar}{r_{jk}^3} \tag{6.3}$$

It is induced by the local magnetic field at spin I_j due to the magnetic moment $\boldsymbol{\mu}$ of a neighboring spin I_k, γ_j and γ_k are the gyromagnetic moments of the spins I_j and I_k and $\mathbf{I_j}$ and $\mathbf{I_k}$ their angular momentum operators, θ is the angle between the static magnetic field $\mathbf{B}_0$ and $\mathbf{r}$ the vector between both spins with its length r.

The dipolar interactions contribute strongly to the signal broadening and are removed by the combined use of magic angle spinning (see MAS section) and radio-frequency irradiation (decoupling). The dipolar couplings among ^{1}H spins are the strongest ones involved in biomolecules. Having a look at their values (table 2) it is obvious why proton spectra are generally much broader and ^{13}C detection is commonly preferred in solid-state NMR, in contrast to the ^{1}H detection usually performed in solutions. For example C-H dipolar couplings in solids can take values of about 20 kHz and are consequently stronger than in liquids. Therefore a much stronger decoupling of 100 kHz or more is required.

Table 2

d_{IS}	1Å	3Å	5Å	7Å
d_{HH}	120	4.44	0.96	0.35
d_{HC}	30	1.11	0.24	0.09
d_{HN}	12	0.44	0.1	0.03
d_{CC}	7.5	0.28	0.06	0.02
d_{CN}	3	0.11	0.02	0.008
d_{NN}	1.2	0.04	0.01	0.003

Table 2 – Dipole-dipole couplings for different distances given in kHz.

Dipolar couplings are used in numerous studies in solids or liquid crystals for distance measurements or the characterization of dynamics. Even in solution they are reintroduced by a partial alignment of the molecules in gels, alcohol, and phages for the measurement of Residual Dipolar Couplings (RDC) [11]. Distances can be directly calculated from the coupling measured in the spectrum over the dipole-dipole coupling constant (equation 6.3).

6.1.2 *Chemical Shift Anisotropy*

Another important contribution to the lineshape is the chemical shift anisotropy. The resonance frequency of an atom is depending on the local magnetic field it is perceiving. Therefore, it is strongly related to its chemical environment as the electrons surrounding the nucleus are shielding it from the static magnetic field. Being most of the time not spherically symmetric, this shielding effect is correlated to the orientation of the molecule regarding

to the external magnetic field. The chemical shift is expressed as a second-rank tensor $\boldsymbol{\sigma}$:

$$\begin{aligned} \mathcal{H}_{CSA} &= \gamma\hbar\mathbf{B}_0\boldsymbol{\sigma}\mathbf{I} \\ &= \gamma \begin{pmatrix} B_{0x} & B_{0y} & B_{0z} \end{pmatrix} \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \begin{pmatrix} I_x \\ I_y \\ I_z \end{pmatrix} \end{aligned} \tag{6.4}$$

In liquids it is averaged by the molecular tumbling to a scalar value, the chemical shift.

If we express the chemical shift tensor of equation 6.4 in its principal axis system (PAS) like shown in figure 45 we obtain its diagonalized form:

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \tag{6.5}$$

The chemical shift tensor actually measured in the static experiment can then be written as a tensor reflecting the variations of the chemical shift depending on the orientation of the molecule. The chemical shift δ_{zz} in a particular orientation to the magnetic field can be expressed as,

$$\delta_{zz} = \lambda_1^2\sigma_{11} + \lambda_2^2\sigma_{22} + \lambda_3^2\sigma_{33} \tag{6.6}$$

where λ_i determines the dependance of the chemical shift to the orientation of the molecule in the static magnetic field B_0.

$$\begin{aligned} \lambda_1 &= \cos\alpha\sin\beta \\ \lambda_2 &= \sin\alpha\sin\beta \\ \lambda_3 &= \cos\beta \end{aligned} \tag{6.7}$$

The Euler angles α and β are relating the PAS of the molecule to the laboratory frame (figure 45) where the z axis is parallel to B_0.

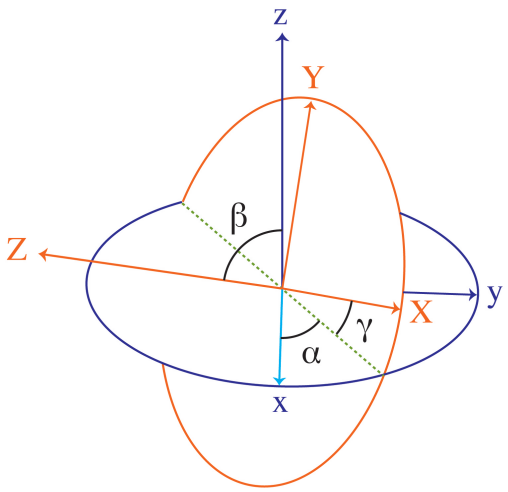


Figure 45 – Representation of the Euler angles α, β and γ relating the PAS system of the chemical shield tensor with its coordinates (X,Y,Z) to the laboratory frame (x,y,z).

6.1.3 *Quadrupolar interaction*

Quadrupolar coupling only occurs in atoms having a spin number higher than $\frac{1}{2}$, thus having more than 2 spin states. The coupling is present even in the absence of an external magnetic field and is due to the environmental electric field gradient (EFG), induced by the electrons of the chemical environment. Therefore highly symmetric molecules do not show quadrupolar interaction.

An approximation of the Hamiltonian for the quadrupolar coupling at high fields can be written as,

$$\mathcal{H}_Q = \frac{3eQ}{4I(2I-1)\hbar}[V_{zz}(3\mathbf{I}_z^2 - \mathbf{I}^2)] \tag{6.8}$$

where eQ is the quadrupolar interaction constant, I the spin number and V_{zz} the z-component of the electric field gradient.

In biomolecular NMR the most important quadrupolar coupled atom is ^{2}H. It can be used to calculate order parameters in solids to characterize dynamics and is often coupled with molecular dynamics (MD) for validation and calibration of the simulations [171].

In liquids the anisotropic contribution is averaged thanks to the overall tumbling of the molecule. This means, the angle θ (figure 46), describing the orientation of the principle axis of the shielding/dipolar tensor with respect to the static magnetic field, is taking rapidly all possible values. Illustrated on the dipolar interaction this means the typical $(3\cos^2\theta - 1)$

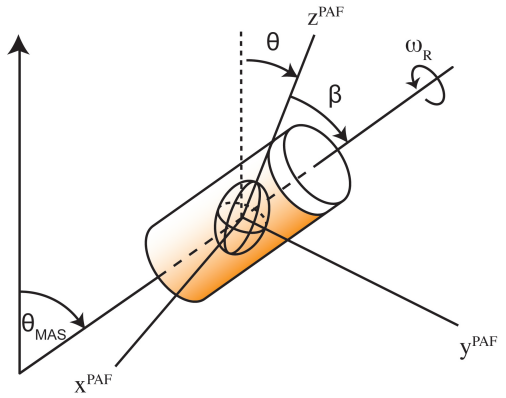


Figure 46 – Principle of the MAS spinning where the sample is rotated at a velocity ω_r at an angle of $\theta_\theta mas$ to the static magnetic field, equivalent to the z axis of the laboratory frame

dependance is averaged to zero on the timescale of the NMR experiment, when the rate of change in the orientation of the molecule is fast compared to the involved interactions. Magic Angle Spinning achieves the same result by mechanically spinning of the sample as shown in figure 46. When the sample is spinning rapidly around an angle θ_{mas} then the angle θ is taking different orientations in regard to the static magnetic field. Therefore, the time average of the orientational dependance of the anisotropic interactions $(3\cos^2\theta - 1)$ can be written as:

$$\langle 3\cos^2\theta - 1\rangle = \frac{1}{2}(3\cos^2\theta_{mas} - 1)(3\cos^2\beta - 1) \tag{6.9}$$

where the angles θ_{mas}, θ and β are defined in figure 46. The angles θ and β are fixed by the chemistry of the molecule, but are taking on all possible orientations in a powder pattern. θ_{mas} is called the magic angle and can be fixed in the experimental setup. If θ_{mas} is set to 54.74° the $(3\cos^2\theta_{mas} - 1)$ term in the time average of $\langle 3\cos^2\theta - 1\rangle$ is zero. Therefore, if the spinning

rate is significantly higher than the anisotropy of the interactions (line-width governed by the static anisotropic interactions), the interactions anisotropy tends to be averaged to zero.

This can be illustrated on the observed chemical shift,

$$\delta = \delta_{iso} + \frac{1}{2}\Delta_{CS}(3\cos^2\theta - 1 + \eta_{CS}\sin^2\theta\cos 2\phi) \tag{6.10}$$

where δ_{iso} is the isotropic chemical shift,

$$\delta_{iso} = \frac{1}{3}(\sigma_{xx}^{PAS} + \sigma_{yy}^{PAS} + \sigma_{zz}^{PAS}) \tag{6.11}$$

Δ_{CS} the chemical shift anisotropy,

$$\Delta_{CS} = -\omega_0(\sigma_{zz}^{PAS} - \sigma_{iso}) \tag{6.12}$$

and η_{CS} the chemical shift asymmetry.

$$\eta_{CS} = \frac{\sigma_y - \sigma_x}{\sigma_z} \tag{6.13}$$

with,

$$\sigma_\alpha = \sigma_{\alpha\alpha}^{PAS} - \sigma_{iso} \tag{6.14}$$

Very fast MAS is averaging the anisotropic tensor to zero and the chemical shift is averaged to its isotropic value. If the MAS speed is not much higher than the chemical shift anisotropy, spinning side-bands are appearing in the spectrum. The MAS averaging is still done over one complete period of the rotation, but the acquisition points are taken with a dwell-time smaller than the rotation period time τ_m. This leads to an modulation of the acquired signal by the MAS rotation frequency ω_r, giving rise to the spinning-sidebands after Fourier transformation. However, Δ_{CS} and η_{CS} can be determined from the spinning side-band pattern giving structural information [115]. Similar analysis can be done by using the anisotropy of dipolar and quadrupolar interactions.

6.3 CROSS-POLARIZATION AND DIRECT EXCITATION (CP-MAS, DD-MAS)

Generally initial magnetization can be created by classic direct excitation with a 90° pulse (DD-MAS) on the nucleus that will be detected as shown in figure 47a. However, the most common nuclei measured in solids are ^{13}C and more rare ^{15}N that have a lower gyromagnetic ratio γ than the more abundant ^{1}H. Longitudinal relaxation constants can also be significantly higher for heteronuclei in solids as in solution, for example ^{13}C T_1 values can reach about 100 s, compared to ^{1}H T_1 time constants that are generally in the order of 3-10 s in solids. If direct ^{13}C excitation is used, the slow relaxation rates force very long recovery delays to allow the polarization to turn back to its thermal equilibrium, thus a transfer from ^{1}H spins to the ^{13}C spin can often lead to a significant sensitivity enhancement. This is usually done by an INEPT transfer of the magnetization in solution and by Cross-Polarization (CP) transfer in solids.

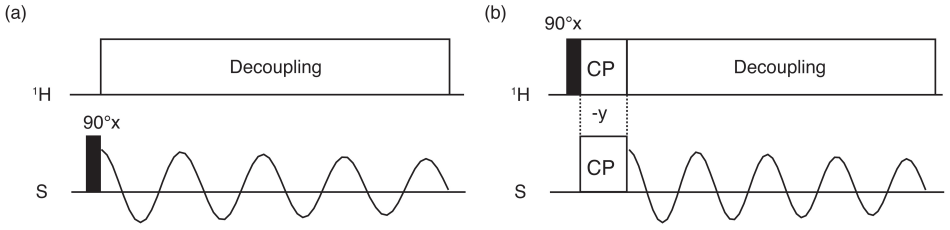


Figure 47 – (a) Classical DD-MAS sequence with direct excitation and homonuclear decoupling while acquisition (b) CP-MAS sequence for polarization transfer from the abundant ^{1}H spins to rare slowly relaxing S spins

In the cross-polarization pulse sequence shown in figure 47b an initial ^{1}H magnetization is created by a 90°_x pulse bringing it along the $-y$ axis. A spin-lock field $B_H(^1H)$ is then applied along the $-y$ axis realized by an on-resonance pulse keeping the ^{1}H spin (to some extent) along the $-y$ axis. Another rf field $B_S(S)$ is applied on the spin S also along the $-y$ axis. If the amplitudes of both rf fields are set to the Hartmann-Hahn conditions [53], a contact is created between both spins, meaning the energy gap in the tilted rotating frame between the spin states α_H and β_H of the ^{1}H spin and between α_S and β_S of spin S are equal.

The Hartmann-Hahn condition in the static case is:

$$\gamma_H B_H(^1H) = \gamma_S B_S(S) \tag{6.15}$$

When MAS is applied they become dependent on the spinning frequency ω_r and have to mach multiples of it to create magnetization transfer:

$$\gamma_H B_H(^1H) - \gamma_S B_S(S) = n\omega_r \, , \, n = \pm 1, \, \pm 2 \tag{6.16}$$

The dipolar coupling between both spins allows then a magnetization transfer from I to S.

The setup of the Hartmann-Hahn condition is done by fixing the amplitude of the contact-pulse of one spin and then varying the pulse amplitude of the other spin until a maximum signal is found. In practice there are two major drawbacks. The biggest problem is that CP dynamics are different for different spin systems, making it depending on the dynamics and the chemical shit of the observed spin. This is a severe problem in quantification studies. Second, in condition of MAS the generally broad Hartmann and Hahn matching curve is broken down into a series of peaks which makes it more difficult to adjust the conditions. Therefore several optimizations of the CP sequence have been proposed. Two mainly used examples are the adiabatic cross-polarization [57], and the ramped amplitude cross-polarization (RAMP-CP) [119, 130].

6.4 CHARACTERIZING DYNAMICS IN THE BACTERIAL CELL WALL

NMR is a very powerful tool to study not only the three-dimensional structure of molecules but also the dynamics they undergo. However, the theoretical bases behind are rather complicated and while a fundamental understanding of the mechanisms in solution has been acquired, the interpretation of dynamics occurring in solids is starting to be developped [44, 45, 98, 113, 154, 168]. However, with the impressive progress made in the last years the number of dynamic studies in solids have increased significantly and various pulse sequences have been developed in this context. A convenience of solid-state NMR is that the dipolar, quadrupolar and CSA interactions are not averaged by the Brownian motion permitting the easy access to slower motions while it is necessary in solution to either reintroduce dipolar couplings by a partial alignment of the molecules (RDCs) or measure chemical exchange rates. However, an important part of the structural and dynamical information included in the anisotropic interactions is averaged by the MAS. Several techniques are used to retrieve this information (figure 48).

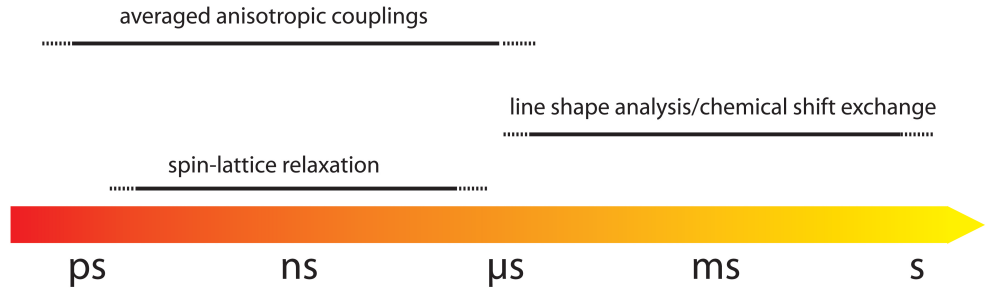


Figure 48 – Schematic representation of different timescales accessible by measuring different solid-state NMR observables. Dynamics can be approximately classified in three different time domains: Fast motions in the order of µs and faster, intermediate or slow motions with motions in the range of µs to ms and finally slow motions at a timescale slower then ms.

One of the major approaches to obtain dynamical information is the lineshape analysis. As MAS is averaging the anisotropic interactions, lowering the spinning speed can be used to retrieve a part of this information contained in the spinning-side bands, appearing when the spinning frequency is lower than the anisotropic interaction. Order parameter describing the impact of local dynamics on the dipolar- and quadrupolar interactions, and the CSA, can be derived from the side-band pattern.

Deuterium is an interesting probe for dynamics in solids. The isotopic enrichment provides a higher selectivity as it allows selective labeling. Being a spin 1 atom, deuterium shows quadrupolar interactions that allow further exploring of the orientation of molecular sites over the well defined direction of the ^{2}H quadrupolar interaction tensor along the C-^{2}H bond [162].

An example of the use of solid state NMR dynamics in the field of bacterial cell adhesion is the work of Rice and Wickham [182] where they were studying structural and dynamical aspects of lipoteichoic acid adhesion on cellulose, peptidoglycan and Tio_2, analyzing ^{31}P CP MAS for changes in the CSA and asymmetry parameters combined with spin-lattice relaxation ^{31} $T_{1\rho}$ in the rotating frame.

6.4.1 *Magnetization transfer as a probe for dynamics in ^{13}C NMR*

Fully hydrated biopolymers like membrane proteins, lipids, polysaccharides or also the teichoic acids are not homogenous in their dynamics. They are often divided in parts with very different types of motions from solid

to solution like. Therefore it is always interesting to perform DD-MAS, CP-MAS, and INEPT magnetization transfers to separate the more solid like portions from the solution like by filtering the very dynamic population by the magnetization transfer.

Regarding equation 6.1 and 6.2 for the dipolar coupling it is obvious that rapid, isotropic motions average the orientational contribution $(3\cos^2\theta - 1)$ to zero and destroy the magnetization transfer. Therefore the CP-MAS sequence can be used to filter highly dynamic contributions out of the spectrum. Another interesting factor is the contact time applied during CP-MAS to allow a transfer between both spins. Varying this parameter can also help to distinguish populations with different dynamical behaviors (CP dynamics). These resonances will still be observed in a DD-MAS sequence like in figure 47a where magnetization is created by direct excitation if the repetition rate is lower than the longitudinal relaxation constant T_1 of the observed spin. On the other hand highly rigid parts can be filtered out using an INEPT transfer since magnetization is canceled out during the transfer period, due to their fast transversal relaxation. In many cases it is useful to compare these different ways of excitation to gain more insights into the dynamical properties of the studied system.

Figure 49 shows on the example of the peptidoglycan of *B. subtilis* 168 how this can be used to quickly characterize dynamical properties by simple recording of 1D spectra. ^{13}C 1D experiments under (a) direct excitation, by initial magnetization transfers using (b) CP, and (c) INEPT are recorded and compared. While the direct excitation gives a global idea of the spectrum as ^{13} T_1 time constants in the peptidoglycan are in the order of 1 s (see chapter 9), and no signal is lost due to the recovery delay of about 3 s, CP is favorable to less dynamic populations and the INEPT transfer to the flexible part of the polymer. Regarding the example of *B. subtilis* 168 it shows clearly that the signals coming from the glycans are much stronger in the CP then in the INEPT experiment while the peptides show an opposite behavior. This points to rigid glycan strands compared to the more flexible peptides.

6.4.2 T_1 *relaxation measurements by inversion recovery*

While the comparison of 1D spectra recorded under different magnetization transfer conditions allows a fast qualitative characterization of dynamics involved, T_1 relaxation measurements allow a quantitative analysis. The most important difference between relaxation studies in solution and in solids is

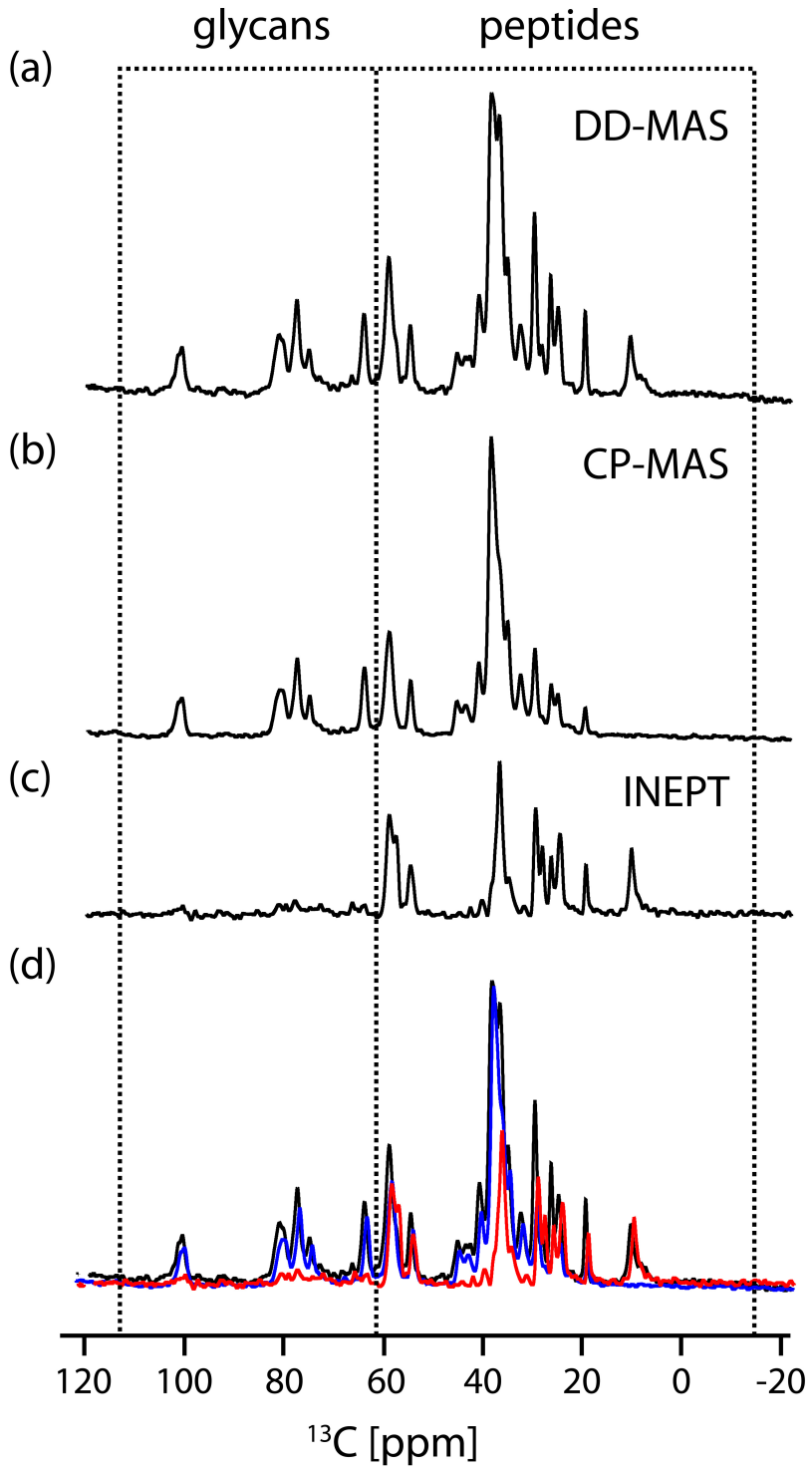


Figure 49 – Spectra recorded on the peptidoglycan of *B. subtilis* 168 using (a) DD-MAS, (b) CP-MAS with a contact time of 800 µs, and (c) INEPT for the creation of the initial polarization. (d) Superposition of the (black) DD-MAS, (blue) CP-MAS, and (red) INEPT experiments.

that the frequently used parameters T_2 and NOE are more complicated to analyse and ambiguous in their interpretation [45, 98, 154, 168]. Nevertheless the T_1 spin-lattice relaxation rates in the laboratory frame and the $T_{1\rho}$ spin-lattice relaxation rates in the rotating frame are easy to measure and give access to motions around the nanosecond range. However, The relaxation in solids is a large and complicated field which is discussed extensively in the literature [113, 168]. Dipolar couplings can lead to spin diffusion in the local environment, thus relaxation constants measured can depend on the local structure of the molecule. Therefore the interpretation of acquired T_1 spin-lattice relaxation time constants has to be done carefully. We limited the interpretation of acquired T_1 spin-lattice relaxation time constants to the comparison between specific sites under different sample conditions.

As in liquid state NMR T_1 longitudinal relaxation time constants can be easily measured by the so called inversion recovery sequence. The pulse sequence is shown in figure 50a. The sequence is simple and consists only out of two radio-frequency pulses separated by an arrayed interval τ. The first pulse is a 180° inversion pulse generating an inversion of the population distribution which is then relaxing back to its thermal distribution while the τ delay. The second pulse is finally creating an observable quantum coherence. Therefore the amplitude of a given peak can be represented as a function of τ.

$$a(\tau) = a(0)(1 - 2e^{-\tau/T_1}) \tag{6.17}$$

Measuring the peak intensity and tracing it with respect to τ gives the typical exponential curve shown in figure 50 and allows to fit a value for T_1.

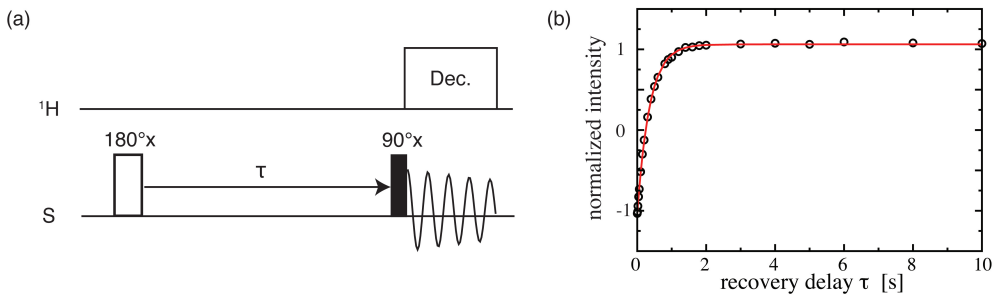


Figure 50 – (a) Inversion recovery pulse sequence (b) Inversion recovery example curve to illustrate exponential dependence of the measured signal to the recovery delay

Because of the overlapping in 1D NMR spectra 2D spectra acquisition is often required. Especially for assignment it is in nearly all the cases inevitable to obtain well resolved 2D spectra (necessitating isotopic labeling). Generally there are two major types of correlation experiments: The first is the classical J-coupling correlation spectroscopy like the Incredible Natural Abundance Double Quantum Transfer Experiment (INADEQUATE) using through bond connectivities [92, 97]. The second type of experiments is using dipolar interactions with the nearest nuclei to create through-space correlations. An example is the Proton Driven Spin Diffusion experiment (PDSD) [66, 165]. In the following section the typical pulse sequences for ^{13}C-^{13}C correlation used in this work are shortly presented. Assignment is largely facilitated by the knowledge about typical chemical shift values for carbon nuclei in different chemical environments. Figure 51 shows a schematic overview over such chemical shift dispersions referenced to TMS. While through-bond spectra give information about covalent connectivities the through-space correlation provides distance constraints. Therefore a combination of both kinds of experiments can furnish precious data on the structure of the studied biopolymer.

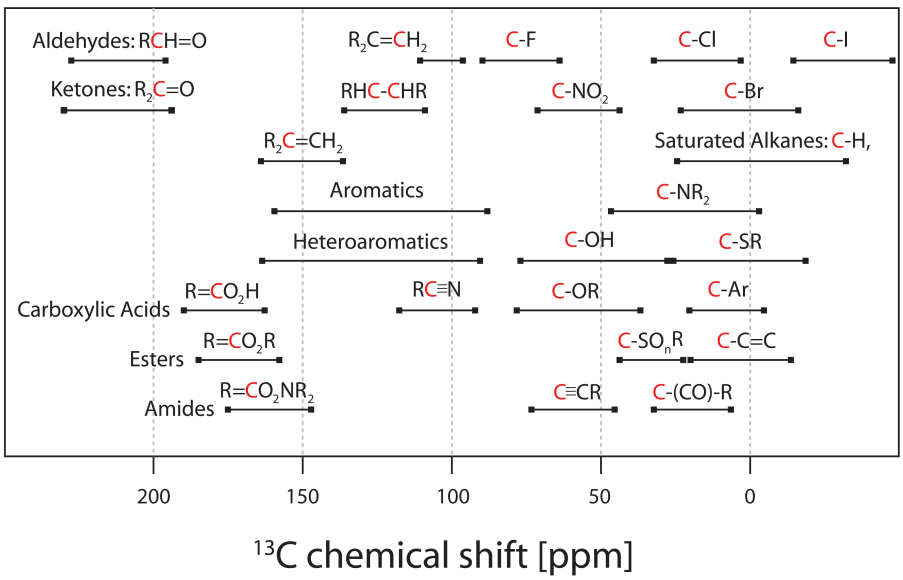


Figure 51 – ^{13}C chemical shift table. All chemical shifts are referenced to TMS. The observed carbon atom is indicated in red

6.5.1 *Through bond connectivities*

In solids the isotropic part of the homonuclear J coupling is generally about 40 times smaller than the dipolar coupling between two nuclei. To give an example the J-coupling between two bond carbon nuclei is about 50Hz, which is in the range of the line-width in the 1D spectra, while the dipolar coupling is at about 2kHz. J couplings have not been commonly used in solid state spectroscopy because much faster magnetization transfer by the strong dipolar interactions, thus dipolar coupling based correlation spectroscopy is generally more sensitive. Nevertheless, it has been shown that J-based magnetization transfer is possible and several pulse sequences have been developed and made through-bond correlation spectroscopy more and more attractive because of its advantages. It allows an easy and often unambiguous assignment of bond carbons, and it is independent of orientation, offset and chemical shift anisotropy. Another advantage in the context of a very dynamic system like the peptidoglycan, where the dipolar interactions are partially averaged, is that through bond connectivities are permitting an even more efficient transfer than the dipolar coupling, especially in mobile parts of the polymer.

Two examples of J-coupling correlation spectroscopy adapted to solid-state NMR are the INADEQUATE [92] and the UC2QF COSY [16]. The former has been developed in liquid state NMR by Ernst and his coworkers [97] and then applied to solid state NMR by Emsley and others [92]. In the standard INADEQUATE double quantum coherence is created before the evolution period and then reconverted to single-quantum coherence for detection. This gives rise to a 2D spectrum where the indirect dimension (ω_1) represents the chemical shift sum of two bond spins which passed the double quantum filter correlated to the chemical shift of the individual spin in the direct dimension (ω_2). The pulse sequence is shown in figure 52a. The advantage is the suppression of diagonal auto-correlation peaks which can cover close cross-peaks. Thus peaks with small chemical shift differences can be easily detected. Nevertheless the classical INADEQUATE sequence has the major drawback that it is creating in- and anti-phase peaks. This results in canceling crosspeaks when the coupling is not resolved. Thus, Emsley and his coworkers proposed a refocused INADEQUATE version of the sequence [92], initially developed for liquid state NMR [97], to overcome this problem. The sequence is shown in figure 52b. The introduced refocusing period is creating only in-phase signals and is therefore far better adapted to solid-state with its broad line-widths.

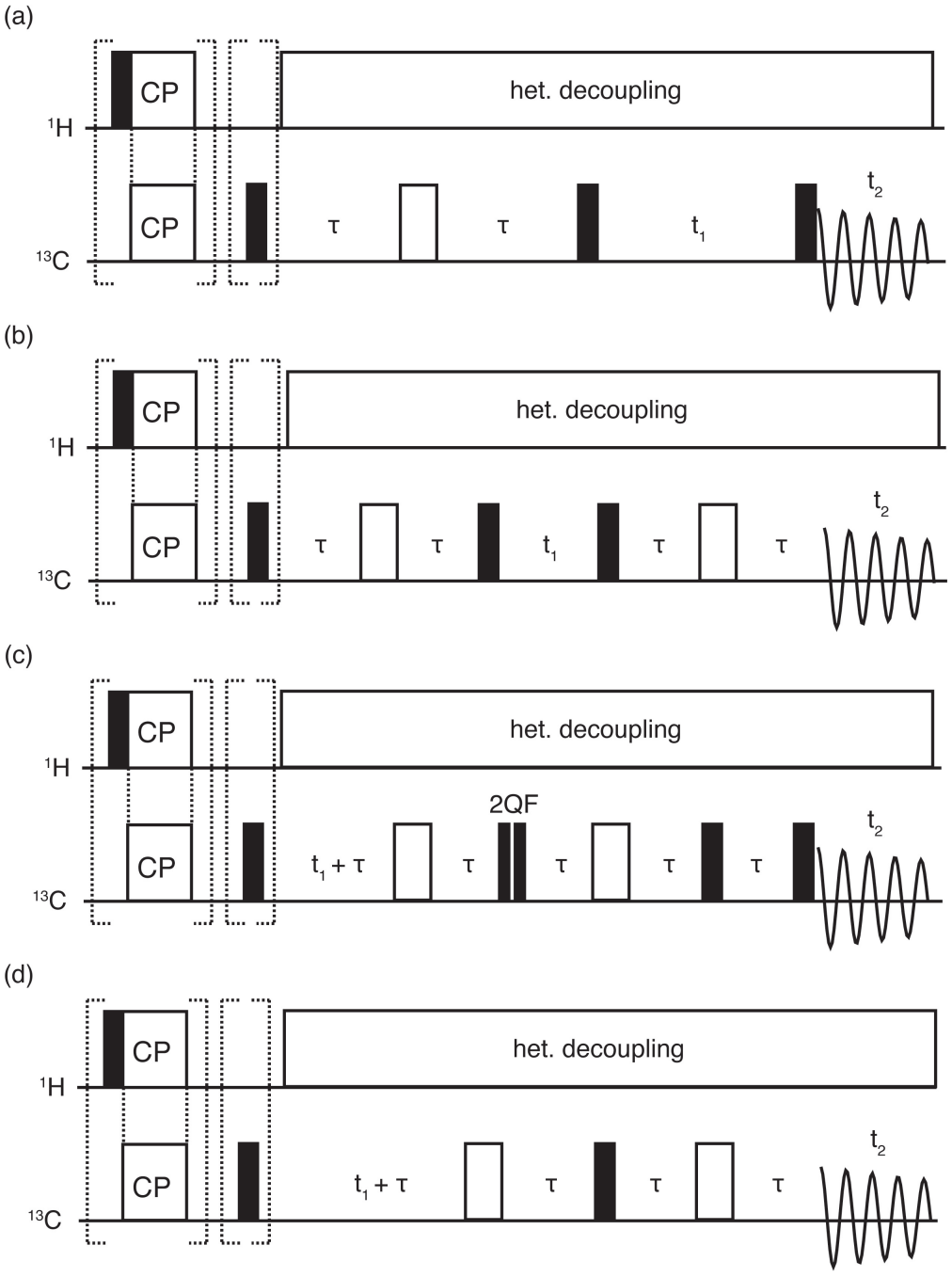


Figure 52 – (a) Pulse sequence for INADEQUATE ^{13}C multiple quantum correlation (b) Pulse sequence for refocused INADEQUATE ^{13}C multiple quantum correlation (c) Pulse sequence for UC2QF COSY (d) Pulse sequence for ^{13}C single quantum correlation

Another alternative for through-bond correlation spectra is the Uniform-sign Cross-peak Double Quantum Filtered Correlation Spectroscopy (UC2QF COSY) [16] also originally developed for solution NMR as double quantum filtered correlation spectroscopy (2QF COSY) and then applied to solid state NMR by Reed and his coworkers in its refocused form also providing in-phase cross peaks. The advantage of the single-quantum experiment is that it is avoiding the faster relaxation of the double quantum part of the INADEQUATE. Figure 52c is showing the published sequence while 52d shows an adapted version which was used for acquisition of the spectra of the peptidoglycans shown in this work. The absence of the double-quantum filter leads to a sensitivity gain. In this work direct excitation was used instead of cross-polarization transfers from the more abundant proton spins. The advantage of the direct excitation in mobile systems like the rather flexible bacterial cell wall is that it allows to observe resonances coming from very dynamical parts of the biomolecule where the dipolar interactions are reduced by the motion. The sensitivity in regions with unfavorable CP dynamics like the CO resonances can also be greatly improved, and peak-intensity comparison between correlation spectra of peptidoglycan alone or in interaction with the protein is not biased. Nevertheless, it introduces a slight signal loss for example for the aliphatic carbons. Figure 53a and b show the correlation pattern for the double and single quantum versions of the experiment. The example spectra shown were taken on the peptidoglycan sample of *E. coli* and demonstrate the typical chemical shift regions for glycans, peptides and carboxyl(carbonyl) correlations to the carbons of the peptide chains. Mainly the single quantum experiment was used for the assignment of the different bacterial species.

6.5.2 *Through space connectivities*

The most commonly used way of coherence transfer in solid state NMR is the dipolar interaction and therefore the trough-space connectivities because of the strong and efficient dipolar couplings. Generally the principles are the same than for the scalar coupling driven transfer. Either double-quantum coherence is created and converted to observable single-quantum coherence or two single-quantum coherences are correlated with cross peaks arising from spins which are close in space.

A classical example is the PDSD polarization transfer published in the beginning of the 80s by Maciel and his coworkers [165]. The experiment was directly adapted from liquid state where it was developed by Ernst to study

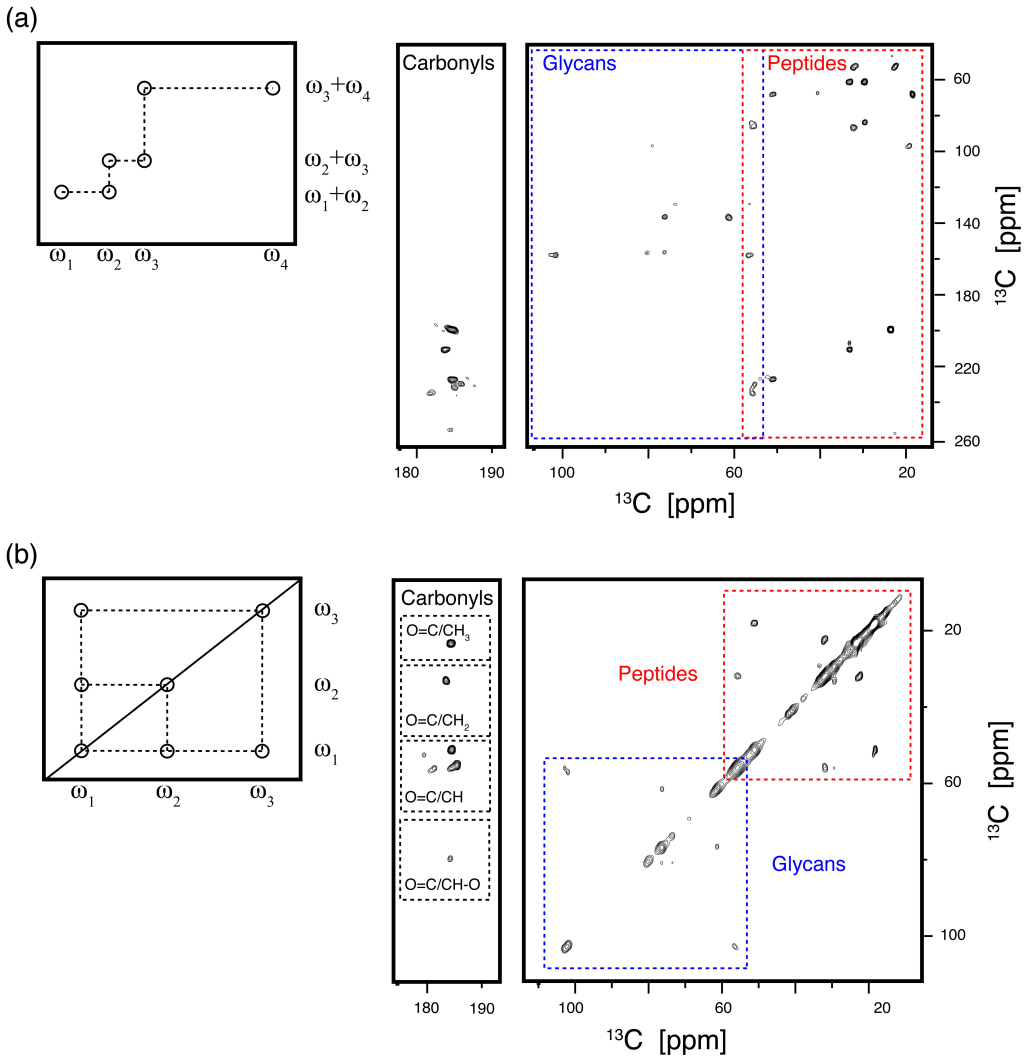


Figure 53 – (a) INADEQUATE ^{13}C double quantum correlation spectrum and (b) SQ ^{13}C correlation spectrum of the peptidoglycan of *Escherichia coli*

exchange and spin diffusion [66]. The sequence is shown in figure 54a. After creation of ^{13}C magnetization in the x-y plane through cross-polarization the spins are labeled while the t_1 evolution period with their characteristic resonance frequency. The polarization transfer is this time created by a mixing period prior to creation of transverse magnetization and detection. Magic Angle Spinning is used to supply the necessary resolution in both dimensions, thus dipolar couplings are averaged to zero and second order term broadening by ^1H spin couplings are needed for the magnetization transfer between the ^{13} spins (proton driven spin diffusion). In consequence, the proton decoupling has to be off during the mixing period to allow spin correlation.

Examples of PDSD correlation spectra on the peptidoglycan of *Escherichia coli*, recorded at different mixing times, are shown in figure 54b-d. The spectra illustrate the impact of mixing time on the cross-peak intensities. The closer the spins are in space the higher the intensity of the cross-peak. With rising mixing times the magnetization is transfered further in space. Therefore dipolar coupling driven correlation spectra give a good idea about spatial distribution, but also on the dynamics of the studied polymer as the dipolar driven magnetization transfer is largely influenced by its dynamics. Flexible parts of the polymer show weaker spin diffusion. On the example of long mixing periods as shown in figure 54d this behavior is visible when comparing the signals from the dynamic peptides to the more rigid glycans.

6.5.3 *Precise distance determination via dipolar couplings*

More precise techniques to measure distance constraints are for example the REDOR [50] and the CODEX experiment [18, 158]. They will briefly be introduced here as they have been applied extensively on freeze dried peptidoglycan samples of *S. aureus*, *B. subtilis*, and *E. faecium* by Schaefer and his coworkers[63, 74–76, 128, 129, 158].

Rotational-echo double resonance (REDOR)

The rotational-echo double resonance (REDOR) sequence was initially developed by Schaefer and his coworkers to measure heteronuclear dipolar couplings under magic angle spinning for distance determination [50]. Under MAS conditions the dipolar-couling is averaged out. Hence the aim of

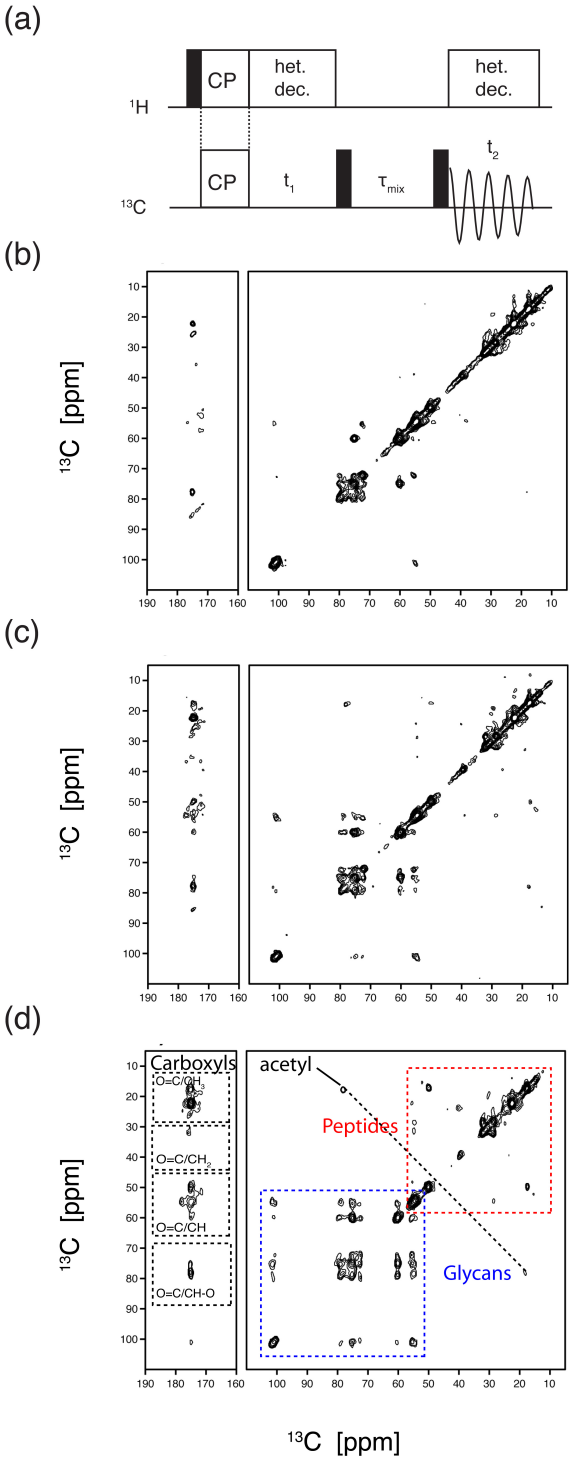


Figure 54 – (a) Pulse sequence for PDSD through space ^{13}C correlation experiments and example spectra recorded on ^{13}C labeled peptidoglycan of *Escherichia coli* with PDSD mixing times of τ_{mix} = 10 ms (b), 30 ms and (c) 100 ms.

the REDOR experiment is to reintroduce dipolar-coupling or better said to prevent it from being averaged by rotor synchronized rf pulses.

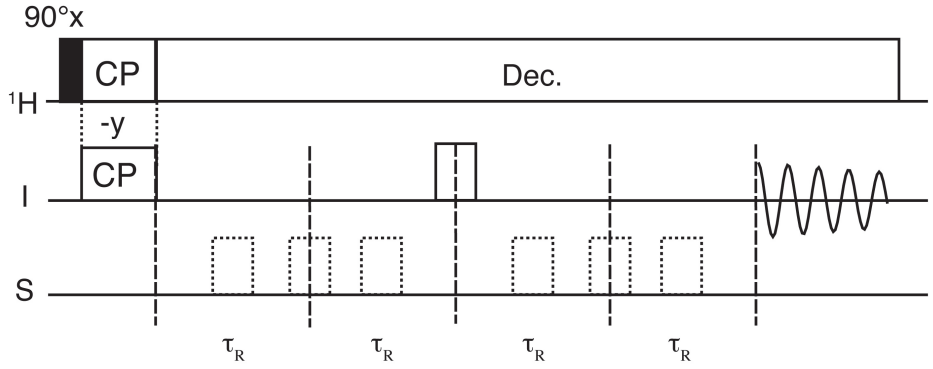


Figure 55 – REDOR pulse sequence. Filled squares correspond to 90° pulses, open squares to 180° pulses. Two experiments are run one with the dotted 180° pulses giving the signal S, and one without giving the reference signal S_0

Under standard conditions transverse magnetization of a given spin I is refocused after every rotor period τ_R. If now a 180° pulse is applied to a coupled spin X every half rotor period the dipolar coupling is not averaged anymore and will induce a dephasing of spin I resulting in an intensity loss in the measured signal S. Comparing the signal S, where the spin I was dephased during time $N\tau_R$, with the reference signal S_0, without the reintroduction of the dipolar coupling by the 180° pulses and therefore only exposed to relaxation effects during the same time, gives access to the dipolar coupling constant D and allows therefore the direct determination of the distance r by fitting the theoretical REDOR equation to the collected data. For the case of an observed nucleus with the spin 1 and a dephasing nucleus with a spin of 1/2 the reduction of the signal is given by:

$$\frac{\Delta S}{S_0} = 1 - \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} \sin\beta \cos(N\tau_R D 2\sqrt{2}\sin 2\beta \sin\alpha)\, d\alpha\, d\beta \qquad (6.18)$$

A drawback of REDOR spectroscopy is that it demands specific spin labels as the flour group of the vancomycin used by Schaefer and his group [74], and very rigid samples are required. In the case of the peptidoglycan it is therefore only feasible on freeze dried samples and not applicable on our hydrated fully labeled peptidoglycan.

Centerband-only detection of exchange (CODEX)

Another way of measuring distances by solid state NMR is the centerband-only detection of exchange introduced by Schmidt-Rohr and his coworkers [18]. Proposed as an experiment to measure dynamics it can also be used for distance determination as shown on the pentaglycine cross-link of *S. aureus* by the group of Schaefer [158]. The pulse sequence represented in figure 56

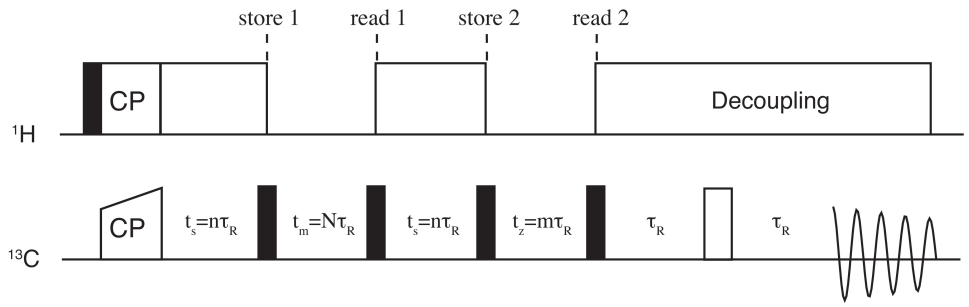


Figure 56 – CODEX pulse sequence. Filled squares are representing 90° pulses open squares are 180° inversion pulses.

is the modified form published by Sharif et al. to measure distances in the peptidoglycan of *Staphylococcus aureus*. If t_m and t_z are in the order as shown in the figure 56 the recorded signal is the CODEX exchange signal S, if they are inverted the recorded signal is the reference S_0. The sum of t_m and t_z is constant and the experiment is therefore relaxation compensated hence the CODEX difference $\Delta S = S_0 - S$ is not including spin-lattice relaxation contributions. All delays are also multiples of the rotor period τ_R and if the parameter n is incremented the two-dimensional CODEX differences can be used to establish the differences in chemical shift of the spins involved in exchange.

In CODEX experiment chemical shift coherences are created during the t_s delay and then stored along the z-direction prior to a long mixing delay t_m after which they are recalled by another 90° pulse. If the chemical shift tensor of the spins remains unchanged during mixing time the spins are refocused, stored during the much shorter t_z delay and then recalled and detected as a Hahn-echo. If the chemical shift is changed while the mixing time, refocusing is not perfect and the resulting signal is altered. Therefore ΔS is nonzero. As anisotropic chemical shift contribution are removed by magic angle spinning and the read and store pulses are rotor synchronized, chemical shift alteration during the mixing time is mainly due to spin-diffusion. By adjusting the mixing time t_m it is therefore possible to allow

spin-diffusion over different distances and to determine distance constraints between different spin systems.

6.6 ^{31}P SPECTROSCOPY

^{31}P NMR is a very suitable technique for biopolymers containing phosphorus nucleus like DNA, RNA, phospholipids or the teichoic acids of the cell wall because of the nuclei spin number of 1/2 and the high sensitivity of ^{31}P. Spectra can be taken directly on non labeled samples because of the 100% natural abundance of the ^{31}P nuclei. Homo-nuclear dipolar interactions resulting into line-broadening are not a problem due to the generally limited number of phosphor atoms compared to abundant atoms like ^{1}H. ^{31}P chemical shifts are also very sensitive to slight chemical changes in the vicinity of the phosphorus atom and to pH and ionic strength. This makes them an ideal probe being in the same time sensitive and natural abundant.

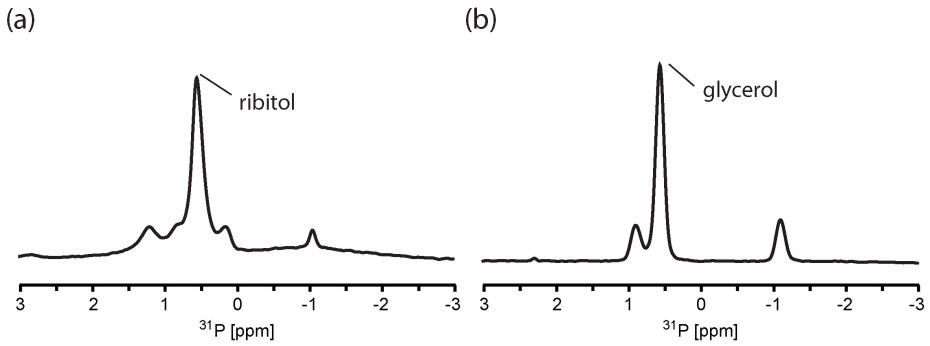


Figure 57 – ^{31}P spectra of (a) complete cells of *B. subtilis* 168 and (b) isolated cell walls of *B. subtilis* 168. Spectra were acquired in an experimental time of about 10 min.

Figure 57 shows two example spectra taken on (a) complete cells and (b) isolated cell walls of *B. subtilis* 168. Both spectra were taken in an experimental time of about 10 min, showing the good sensitivity of the experiment. Also, only signals coming from the teichoic acids are observed in the bacterial cell wall as they are the only molecules containing phosphor groups. This allows to analyze structural and dynamical properties of the polyphosphate polymer directly in the peptidoglycan network with minimum overlaps in the spectra obtained.

Part V

INSIGHTS IN THE BACTERIAL CELL WALL BY SOLID STATE NMR

7

CHARACTERIZATION OF PEPTIDOGLYCAN STRUCTURE AND PROTEIN INTERACTIONS

7.1 CHARACTERIZATION OF THE STRUCTURE AND DYNAMICS OF THE PEPTIDOGLYCAN FROM DIFFERENT GRAM BACTERIA

Despite decades of research, the structure of bacterial cell walls has remained a challenging topic due to the enormous size (up to $3x10^6$ kDa for *E. coli*) and complex heterogeneous composition of these biomolecules precluding structural determination by established methods such as X-ray crystallography. Although the chemical composition of the peptidoglycan and teichoic acids of a large number of bacteria has been determined, very little is known about their three-dimensional structure. Biomolecular solid-state nuclear magnetic resonance spectroscopy (ssNMR) has made huge progress in the last years, becoming a promising method for the structural characterization of non-crystalline biopolymers [112]. Currently only a few ssNMR studies have been performed on specifically labeled bacterial cell-wall samples. Structural information on interactions between antibiotics and the peptide cross-links of the peptidoglycan have been obtained using ^{13}C REDOR experiments [74, 76], and adhesion of lyophilized LTA samples to various surfaces have been studied by ^{31}P solid-state NMR [137]. In this work, we show for the first time that solid-state NMR spectroscopy can successfully be applied to intact hydrated and fully labeled sacculi of *E. coli*, *B. subtilis*, *S. aureus* and *S. pneumoniae*. A combination of solid-state ^{13}C-^{13}C correlation experiments proved to provide key structural information.

We show that it is possible to obtain surprisingly well-resolved solid-state ^{13}C-^{13}C NMR correlation spectra from hydrated cell wall samples of a range of different species. Based on first results on the sacculi from E. coli [71] we were also able to determine solid-state NMR data from cell walls and even intact cells of Gram-positive species. These data provide novel and

important insights into the flexibility and molecular arrangements of the two major cell wall components in Gram-positive bacteria, peptidoglycan and wall teichoic acid (see chapter "characterization of the wall teichoic acids"). In addition, this work shows that solid-state NMR spectroscopy is a powerful tool to study heterogeneous biological macromolecules.

7.1.1 *Characterization of the peptidoglycan from the Gram-negative bacterium E. coli*

^{13}C-trough bond and through space correlation experiments were performed, and spectra of astonishing good resolution were obtained, considering the size and noncrystallinity of the sample. Figure 58b illustrates the high spectral quality obtained using a through-bond correlation SQ/SQ correlation experiment. The spectrum was acquired using the sequence in Figure 69, which is closely related to previously published experiments [16, 92]. Interestingly, the signals observed for the glycan part present symmetric line shapes, without any distortion due to chemical shift inhomogeneity, as has been observed, for example, in cellulose [142]. This suggests a structural inhomogeneity which is small enough to be averaged out by the amplitude-restricted molecular dynamics present in the cross-linked polymer.

The flexibility of the peptidoglycan polymer can be evaluated from the transfer efficiency observed in proton-driven spin diffusion experiments (PDSD) [165], as an increased mobility yields reduced dipolar couplings. At short contact times ($\approx$ 2 ms), very few cross peaks are detected, indicating the presence of significant dynamics along the entire polymer network. If the contact time is increased to about 10 ms, correlations between spatially close carbons in the sugar rings are mainly observed. This suggests that the glycan strands are more rigid than the peptide chains, for which only few correlations are detected even for mixing times as long as 100 ms (Figure 59). The dynamics of the polymer network contribute to the high resolution observed in the NMR spectra.

Essentially complete resonance assignment of peptidoglycan of *E. coli* has been obtained (Figure 58b). The glutamate and the lactic acid spin systems were unambiguously assigned. Interestingly, two sets of cross peaks were observed for DAP, which can be explained by the peptide cross-linking biodiversity (D-Ala-D-m-DAP and D-DAP-L-DAP) in *E. coli* [173]. Resonances of the two sugar units could not be unambiguously assigned. However, two resolved peaks were observed for the C-1, C-2, and C-3 sites of the

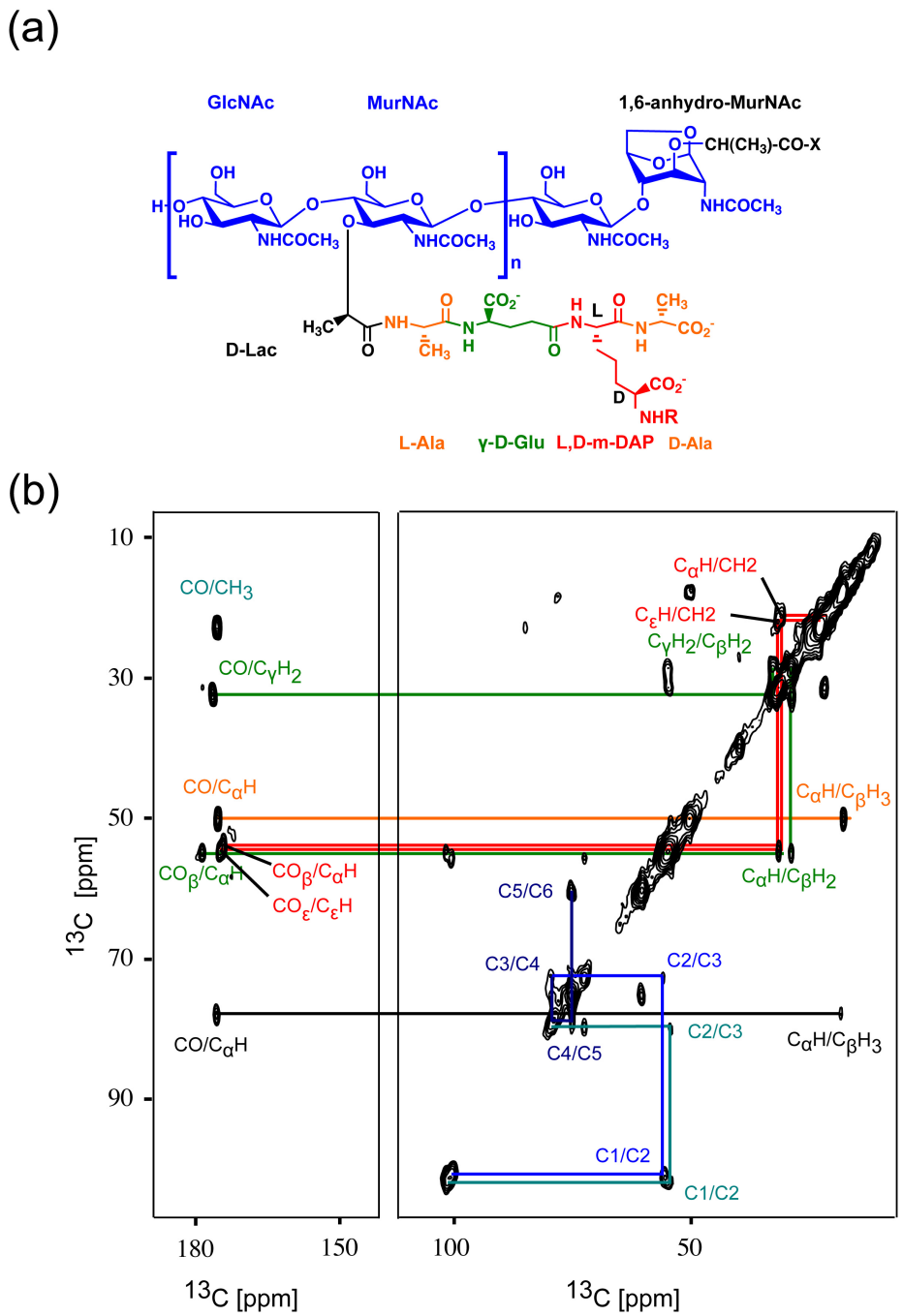


Figure 58 – (a) Chemical structure of E. coli peptidoglycan. The letter R stands for a hydrogen atom or the peptide stem of another peptidoglycan chain; X stands for a peptide stem. (b) Through-bond ^{13}C-correlation spectrum of ^{13}C-labeled peptidoglycan, measured on a 400 MHz (^{1}H frequency) spectrometer. The resonance assignment follows the color code defined in panel (a) for the different spin systems: shades of blue for glycan, black for lactic acid, orange for alanines, green for glutamate, and red for diaminopimelic acid.

sugars. Additional resonances that could account for the small amount of 1,6-anhydro-MurNAc (3-6%) were not detected. The L- and D-alanine could not be unambiguously assigned.

7.1.2 *Characterization of the peptidoglycan from the Gram-positive bacteria B. subtillis 168 and S. aureus*

Cell wall samples from *B. subtilis* and *S. aureus* were uniformly ^{13}C and ^{15}N-isotopically labeled and prepared similarly to the method described for *E. coli* [71]. This procedure retained the intact cell envelope (sacculi) formed by the peptidoglycan and covalently linked WTAs, while LTAs were eliminated through the SDS treatment of the bacterial cells in the isolation steps [10, 110]. Additional treatment of the cell wall with hydrofluoric acid removed the WTA by hydrolysis of the phosphate group attached to the position 6 of the N-acetylmuramic acid. Thus, hydrated uniformly ^{13}C- and ^{15}N-labeled peptidoglycan samples, with and without attached WTAs, were obtained from *B. subtilis* and *S. aureus*.

SQ/SQ ^{13}C-^{13}C J-correlation experiments were recorded on cell wall and peptidoglycan samples from *B. subtilis* and *S. aureus* and compared to the spectra previously recorded on *E. coli*. Within the same species, spectra only differed by a few distinct peaks that could consequently be unambiguously assigned to teichoic acids, while the chemical shift of the resonances from the peptidoglycan remained unaffected. This comparison confirmed that the peptidoglycan structure is not altered by the removal of the teichoic acids in either strain.

^{13}C spectra of the peptidoglycan of Gram-positive bacteria were further compared with the one measured on the *E. coli* Gram-negative bacterial model. Figure 60a details this comparison for the carbonyl-to-aliphatic region of the ^{13}C-^{13}C correlation spectra of *E. coli*, *B. subtilis* and *S. aureus*, and reveals a very limited number of differences. In the case of *B. subtilis*, additional peaks are observed in the CO/C_α or CO/C_γ or CO/C_ϵ chemical shift regions corresponding to glutamic acid and meso-A_2pm residues. This is consistent with the formerly described partial amidation of meso-A_2pm in the peptidoglycan of *B. subtilis* [3, 177]. For *S. aureus*, the most significant differences compared to *E. coli* arise from the intense signal assigned to the glycine residues of the penta-glycine inter-peptide bridge, from the presence of L-lysine instead of meso-A_2pm in the peptide strands and from

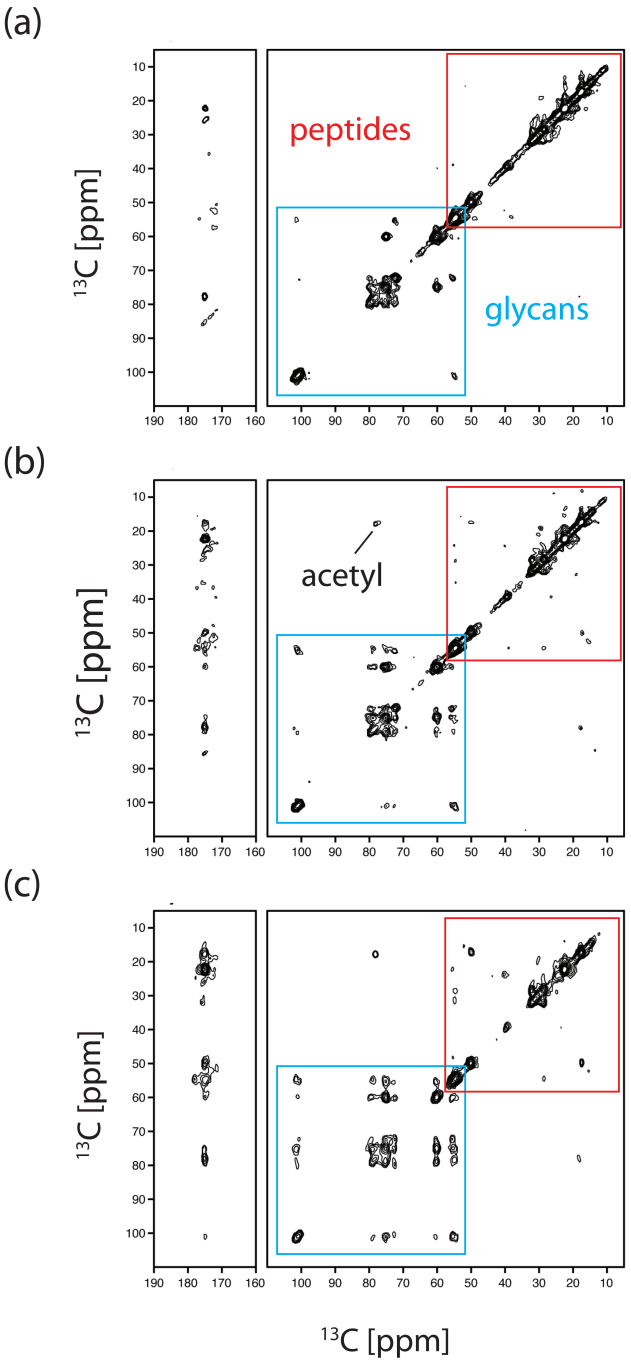


Figure 59 – (a-c) PDSD through-space ^{13}C correlation experiments of ^{13}C labeled peptidoclycan measured with a PDSD mixing time of τ_{mix} = 10 ms (a), 30 ms (b), and 100 ms (c). Initial ^{13}C magnetization was obtained by adiabatic cross polarization [4]. The indirect acquisition time t_1 was set to 9.5 ms, 10 ms, and 6 ms for (a), (b), and (c), respectively. States quadrature dectection was applied in the indirect detection dimension.

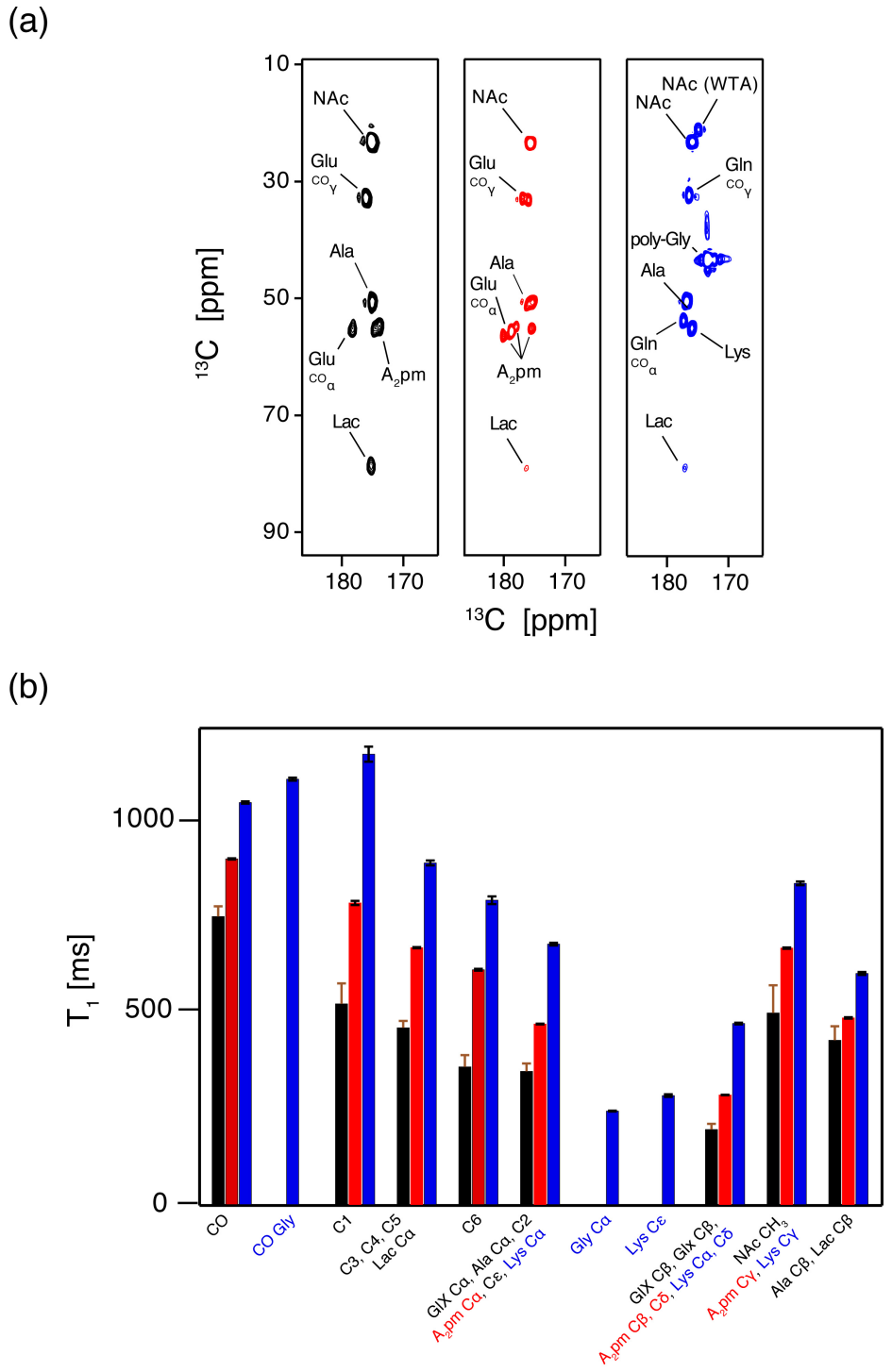


Figure 60 – Comparison of (a) the carbonyl-to-aliphatic region of the ^{13}C-^{13}C J-correlation experiment collected on the 400MHz Bruker spectrometer, and of (b) the T_1 longitudinal relaxation time constants collected on peptidoglycan samples from *E. coli* (black), cell walls of *B. subtilis* 168 (red) and *S. aureus* (blue). T_1 relaxation time constants reported in part (b) were measured on WTA-free samples.

the chemical shift variations of the glutamic acid CO-to-C_α correlations due to significant amidation of the Glu residue into Gln.

To compare the overall and the local flexibility of the peptidoglycan network of Gram-negative and Gram-positive bacteria, carbon longitudinal relaxation time-constants T_1 were measured for peptidoglycans from *B. subtilis* and *S. aureus* using inversion-recovery experiments (Fig. 60b). These relaxation time-constants were then compared to the T_1s obtained for peptidoglycan isolated from *E. coli* [71]. These measurements show a homogeneous increase of the T_1 values for all resonances of the peptidoglycan of *B. subtilis* compared to *E. coli* and of *S. aureus* compared to *B. subtilis*. As verified by measuring the T_1 relaxation rates for different temperatures showing an decrease in flexibility for increasing relaxation time constants (see annex D) in the case of the peptidoglycan. This points out a global decrease of the molecular flexibility of the peptidoglycan in the following order: *E. coli* $>$ *B. subtilis* $>$ *S. aureus*. Hence the peptidoglycan of *S. aureus* is the most rigid of the three peptidoglycans studied.

7.1.3 *Discussion*

Spectra obtained from thick *B. subtilis* peptidoglycan sacculi were highly similar to spectra from thin *E. coli* sacculi (Fig. 60). The similarity of the NMR signals shows that the local organization of the two peptidoglycan samples is similar and is consistent with the related chemical composition of peptidoglycan in both species. Considering the distance between glycan strands which is greater than 30 Å (based on the length of the stem-peptides), we expect that the ^{13}C chemical shifts will be mainly influenced by the local structure inside one glycan or peptide chain rather than by the presence of neighboring chains. Thus, as observed in this experiment, the presence of multiple peptidoglycan layers in Gram-positive species should not affect the chemical shifts of the ^{13}C resonances.

Peptidoglycan from *E. coli* and *B. subtilis* differ mainly by their thickness ($\approx$ 5 nm and $\approx$ 30 nm, respectively) and their average glycan strand length ($\approx$ 30 and more than $\approx$ 500 disaccharide units, respectively). The peptidoglycan of *S. aureus* is characterized by very short glycan strands (in average 5-10 disaccharide units) and an unusually high degree of peptide cross-links, with the presence of multimeric peptides (up to 15-mers) (Fig. 61). Our measurements of carbon longitudinal relaxation time-constants T_1 (Fig. 60b) point out a global decrease of the molecular flexibility of the peptidoglycan

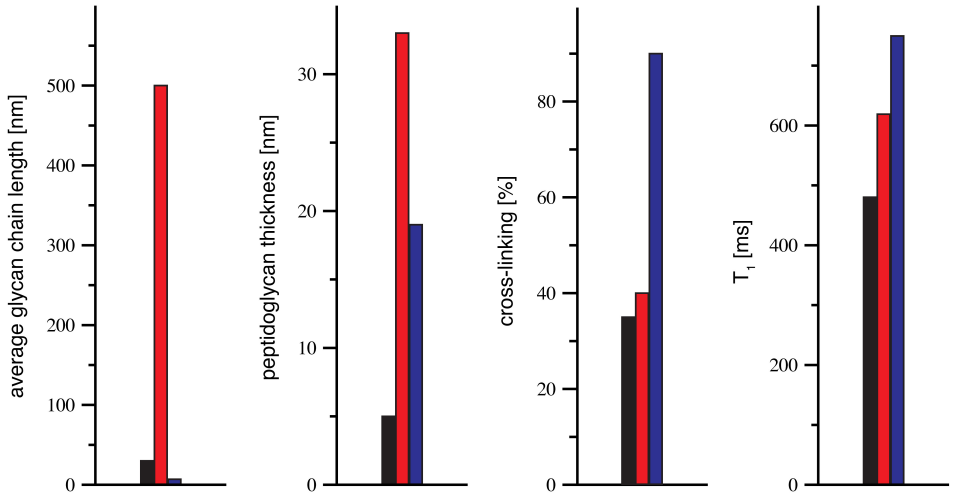


Figure 61 – Variation of peptidoglycan global structural parameters within three bacteria: *E. coli* (black), *B. subtilis* (red) and *S. aureus* (blue). T_1 relaxation time constants correspond to the average values measured on all ^{13}C nuclei.

in the following order: *E. coli* > *B. subtilis* > *S. aureus*, the peptidoglycan of *S. aureus* being the most rigid of the three peptidoglycans studied in this paper. This finding is surprising and maybe counterintuitive. Indeed the glycan strands, which are the more rigid part of the peptidoglycan, are shorter in *S. aureus*, and the flexible part built by the peptide bridges is supposed to have an enhanced mobility due to the presence of the pentaglycines. Obviously, the particularly high degree of peptide cross-linkage over-compensates for the shortness of the glycan strands and the presence of the peptide bridges and causes major decrease in overall flexibility of the *S. aureus* peptidoglycan (Fig. 61). Remarkably, comparison of the different ^{13}C resonances of the peptide stem revealed a relatively high local flexibility for the glycine residues, which, however, is not propagated to the rest of the peptide chain that remains more rigid in *S. aureus* than in *B. subtilis* or *E. coli*, where there is no extended bridging between the crosslinks. This finding shows that the presence of the interpeptide bridge cannot enhance flexibility in peptidoglycan with a high degree of peptide cross-linkage.

7.2 CHARACTERIZATION OF PROTEIN-PEPTIDOGLYCAN INTERACTIONS BY SOLID STATE NMR SPECTROSCOPY

It has been shown in the last section that it is possible to study purified, intact peptidoglycan of *E. coli*, *B. subtilis*, and *S. aureus*. Further, we demonstrated

that Solid State NMR is a powerful tool to characterize their dynamics and chemical properties. Regarding the importance of the biosynthesis of the peptidoglycan in context of antibiotic treatments and also its implication in the response of the innate immune system, it is crucial to get more informations about protein-peptidoglycan interactions. In liquid state NMR the size of the interacting protein is limited and the same size limitation allows interaction studies only in the context of relatively small peptidoglycan fragments. Unfortunately, many proteins involved in this processes are not or weakly interacting with digested fragments of the peptidoglycan and alternatives have to be found for getting structural information about their interactions. The concept of flexibility can also be exploited to characterize peptidoglycan-protein interaction on the entire peptidoglycan. YajG is a lipoprotein where it has been shown that it is interacting with the peptidoglycan while its role stays unknown [12]. As it is a medium sized protein (about 25kDa) with a good solubility, it is very suited as a model protein for testing protein-peptidoglycan interactions. The results presented in this section show that it is possible to characterize its interaction on the peptidoglycan using the techniques presented in the previous section and to obtain structural information about its binding on the peptidoglycan of *E. coli*.

7.2.1 *T_1 relaxation measurements combined with through-bond ^{13}C correlation experiments reveal YajG binding to the peptidoglycan*

With the obtained assignment for the peptidoglycan of *E. coli*, a sufficient base is provided to start screening for interaction sites in intact peptidoglycan sacculi. To study peptidoglycan-protein interaction, sacculi were as described in the previous section incubated with YajG using 50 µL of the previously studied sacculi preparation, incubated for 2 h at 37 °C with 1 mL of a 800 µM protein solution prepared in 50 mM MES at pH 6 (figure 71). After incubation, the sample was centrifuged in the solid-state NMR rotor and a throughbond ^{13}C-correlation experiment was collected and compared to a reference spectrum recorded on a pure peptidoglycan sample conditioned with the same buffer.

No chemical shift changes were observed upon interaction. However, cross-peak intensities were modified significantly and in a nonuniform manner. Figure 62a shows the ratio between the normalized cross-peak intensities in the spectrum of sacculi alone and in interaction with the protein. Glycan and lactic acid sites are clearly altered by the presence of the protein, while

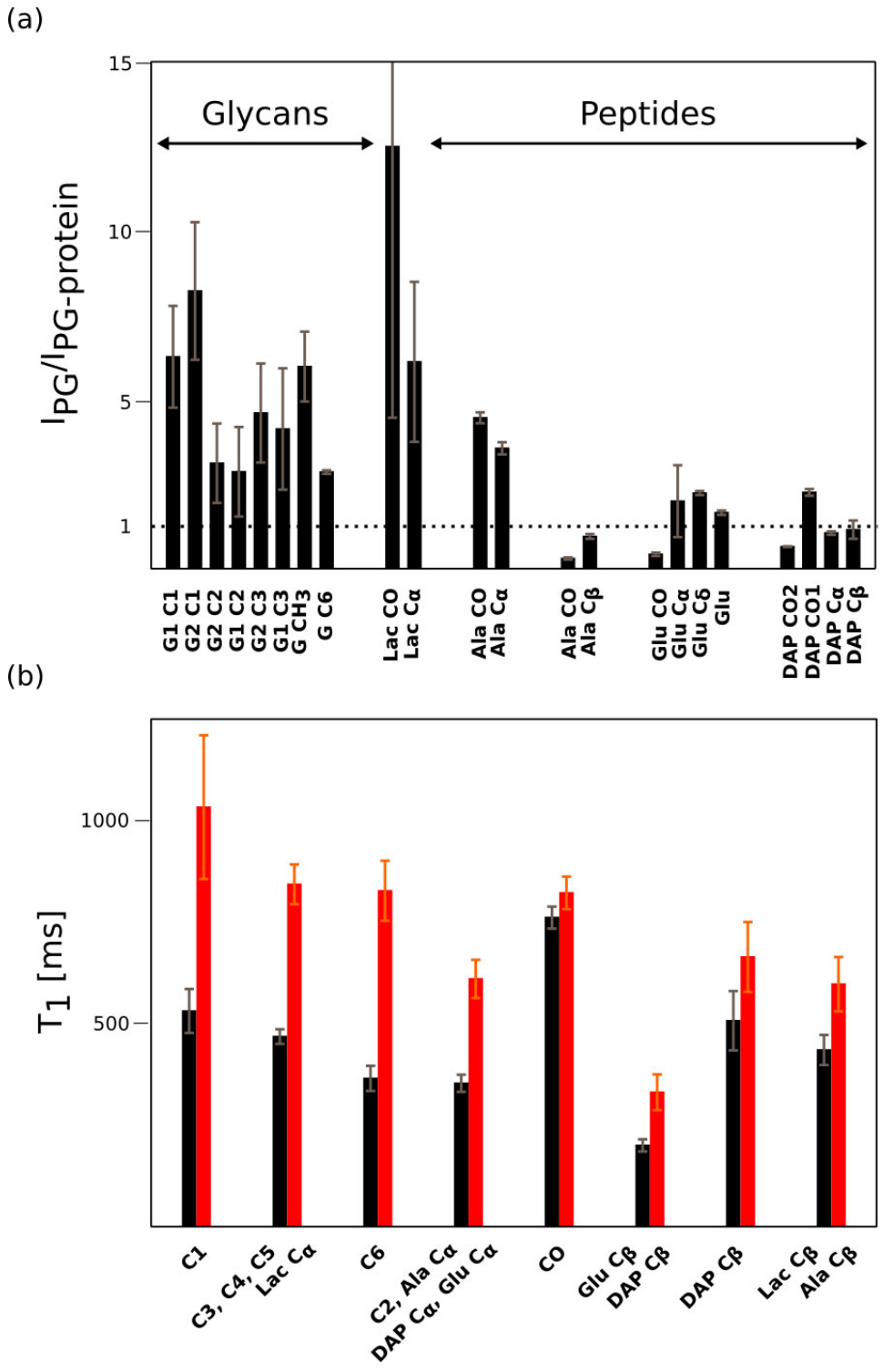


Figure 62 – (a) Ratios of cross-peak intensity in through-bond ^{13}C-correlation spectra of E. coli sacculi alone and in interaction. (b) T_1 relaxation time values fitted from ^{13}C inversion-recovery experiments collected on peptidoglycan alone (black) and in interaction with the protein (red).

peptide cross-peaks remain nearly unchanged. To confirm the glycans as the preferential interaction site, ^{13}C longitudinal relaxation times (T_1) were measured on both samples. Results, presented in Figure 62b, show an overall increase in the ^{13}C T1 relaxation times observed for the protein-interacting sacculi, that is, clearly more pronounced for the glycan resonances. Thus, the interaction with the protein induces an overall decrease in the mobility of the peptidoglycan network but targets primarily the glycan chains. It has to be mentioned that modifications of the T_1 relaxation time constants can also be introduced by spin diffusion through the close by protons of the interacting protein.

7.3 CONCLUSION

In conclusion, we have shown that intact peptidoglycan sacculi of the Gram-negative bacterium *E. coli* give high-quality solid-state NMR spectra, allowing atom-resolved investigation of the peptidoglycan structure and dynamics. We demonstrated that organization and flexibility of different bacterial cell walls can be analyzed by ^{13}C solid-state NMR. By comparing the dynamic behavior of these nuclei, local flexibility can be analyzed and correlated to other physical properties. This strategy will permit the characterization of other specific cell wall properties, for example in the presence of an antibiotic stress. Furthermore, this technique has the potential to elucidate the mechanisms of the interactions between peptidoglycan and proteins such as Penicillin-binding proteins or proteins of the innate immunity [49, 114].

The spectra reflect the small heterogeneity of the peptidoglycan fragments constituting the sacculus, the synthesis and remodeling of which involves a large set of cellular proteins. We were able to observe the interaction of the sacculi with one of these proteins. This contribution demonstrates that solid-state NMR can be a powerful tool to screen for cell-wall interacting proteins.

8

CHARACTERIZATION OF THE STRUCTURE AND METAL BINDING IN THE COMPLETE BACTERIAL CELL WALL OF GRAM-POSITIVE BACTERIA

8.1 INTRODUCTION

We have shown in the last chapter that biomolecular solid state NMR is well suited to study structural aspects of the peptidoglycan. Here we will show that combining the previously presented ^{13}C correlation experiments with ^{31}P NMR spectroscopy is a powerful tool to study complete cell walls of Gram-positive bacteria containing both peptidoglycan and teichoic acids. As pointed out in the introducing chapter to the bacterial cell wall, the polyionic network formed by the teichoic acids has been proposed for diverse roles. The best known is the regulation of ion homeostasis and trafficking of ions in and out of the cell. As until now most NMR studies of the teichoic acids have been done on isolated lipo-teichoic acids (LTAs), we show that our method can be expanded and wall teichoic acids can be directly studied in their natural environment attached to the peptidoglycan. In the aim to gain more insight in the influence of teichoic acids on the structural properties of the peptidoglycan and on the regulation of metal ion homeostasis, we have recorded ^{13}C-^{13}C correlation spectra on intact cell walls for resonance assignment and comparison with spectra recorded on isolated peptidoglycan. Further we have compared ^{31}P NMR spectra on different Gram-positive species to characterize and differentiate wall teicoic acids.

8.2 CHARACTERIZATION OF THE COMPLETE CELL WALL BY SOLID STATE NMR - WALL TEICHOIC ACIDS

Resonances of WTAs have been readily identified by comparing ^{13}C-^{13}C correlation spectra recorded on sacculi with and without WTAs. Using a

combination of characteristic correlation patterns and key chemical shifts, individual resonances of WTAs could be unambiguously assigned (see annex). Their chemical shifts are comparable to ^{13}C data published on extracted lipoteichoic acids or cell wall polysaccharide fragments [80] studied by liquid-state NMR. Carbon spectra of these samples revealed that the WTAs signals could be divided into two fractions. The first had such a long apparent transversal relaxation time that one-bond ^{13}C-^{13}C and ^{1}H-^{13}C J-couplings were completely resolved when using a long acquisition time (50 ms) and no proton decoupling. For them dipolar couplings were completely averaged out as a result of extremely high dynamics of the chains. The second fraction of WTA signals appeared in ^{1}H to ^{13}C cross-polarization transfer experiments implying the presence of other more rigid parts of WTAs. In the case of *S. pneumoniae* a specific labeling of the teichoic acids has been obtained by growing them in a rich C+Y medium [60] and adding ^{13}C labeled glucose. The spectrum shown in figure 63 points out the very well resolved glycan region and its assignment. Detailed assignments of all the bacterial species studied are given in the annex.

To characterize a possible alteration of the WTAs due to the preparation of the cell walls, solid-state ^{31}P 1D experiments were acquired on intact bacterial cells and on the purified sacculi in identical experimental conditions. In fact ^{31}P chemical shifts are very sensitive to chemical substitutions in the vicinity of the phosphorus atom and to a slightly lesser extent to pH and/or ionic strength. The presence of ^{31}P resonances at similar chemical shifts in both bacterial and isolated cell wall samples (Fig. 64) shows that the chemical structure of WTAs is unaffected by the treatment used for cell wall purification, and allows the identification of WTA-specific resonances in complete bacteria. This showed that intact WTAs can be directly studied by solid-state NMR using intact bacterial cells or cell wall preparations. Hence differences in the properties of WTAs can be readily observed using ^{31}P NMR on complete bacteria without the need for cell wall isolation.

The inter-species comparison of ^{31}P spectra from isolated cell wall samples (shown in Fig. 65a-d) emphasizes the potential of solid-state NMR in the differentiation of WTA characteristics. By comparison with data published for LTAs that bear similar phosphate groups to WTAs, partial assignment of the WTAs spectra could be obtained. Common to all WTA spectra was a subgroup of small peaks corresponding to the linkage units of the WTA and one major population representing the long repetitive chain. As shown for *Streptococcus pneumoniae* [77], the two resonances between 0 and -1 ppm (Fig. 65c) can be assigned to the phosphocholine groups and the

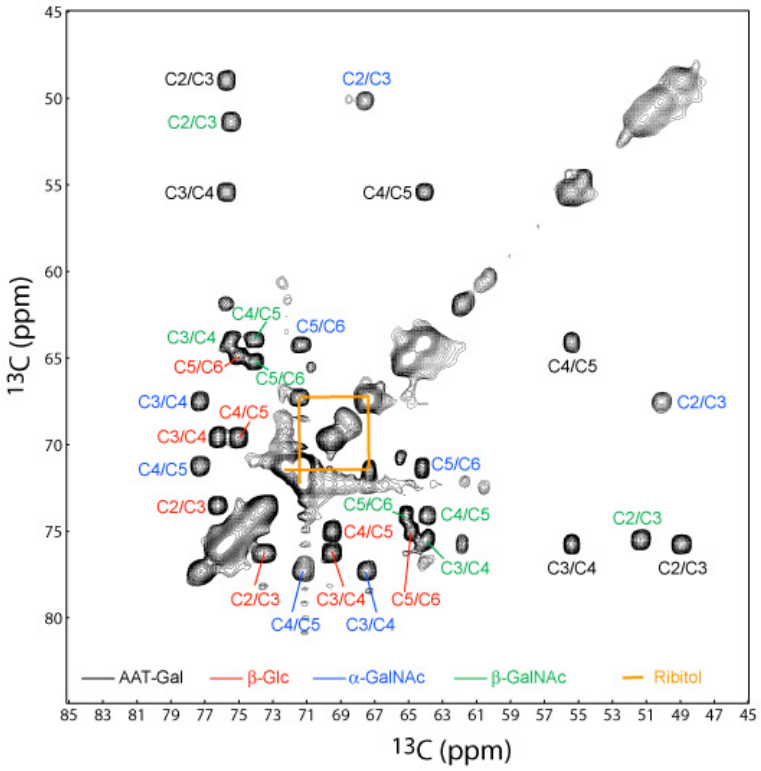


Figure 63 – ^{13}C-^{13}C through bond correlation spectrum of *S. pneumoniae* recorded on a cell wall sample with specifically labeled teichoic acids.

resonance at 1.3 ppm to the ribitolphosphate group of WTA. Assignment of the ribitolphosphate is confirmed by the resonances observed at 1.4 ppm on the spectrum of *B. subtilis* W23, for which WTA repeats are formed by ribitol (Fig. 65b). Glycerolphosphate can be assigned at lower field (0.6 ppm), based on the spectrum of *B. subtilis* 168 (Fig. 65a) in which the main WTA repeating unit is formed by a glycerol instead of ribitol. The remaining less intense peaks in the later spectrum could be assigned to the phosphate group linking the position 6 of the peptidoglycan MurNAc sugar to the GlcNAc of the WTA linkage unit and to the first glycerolphosphate attached to the ManNAc of the linkage unit. Assignment of unsubstituted glycerol in isolated LTA has also been reported in literature, and it was shown that the glycerol OH-to-O-alanyl or OH-to-O-glucosyl substitution increases the phosphorus chemical-shift shielding by approximately 0.2 to 0.7 ppm [6]. A similar effect on ribitolphosphate groups can be postulated and would explain the overlapping peaks observed for *S. aureus* (between

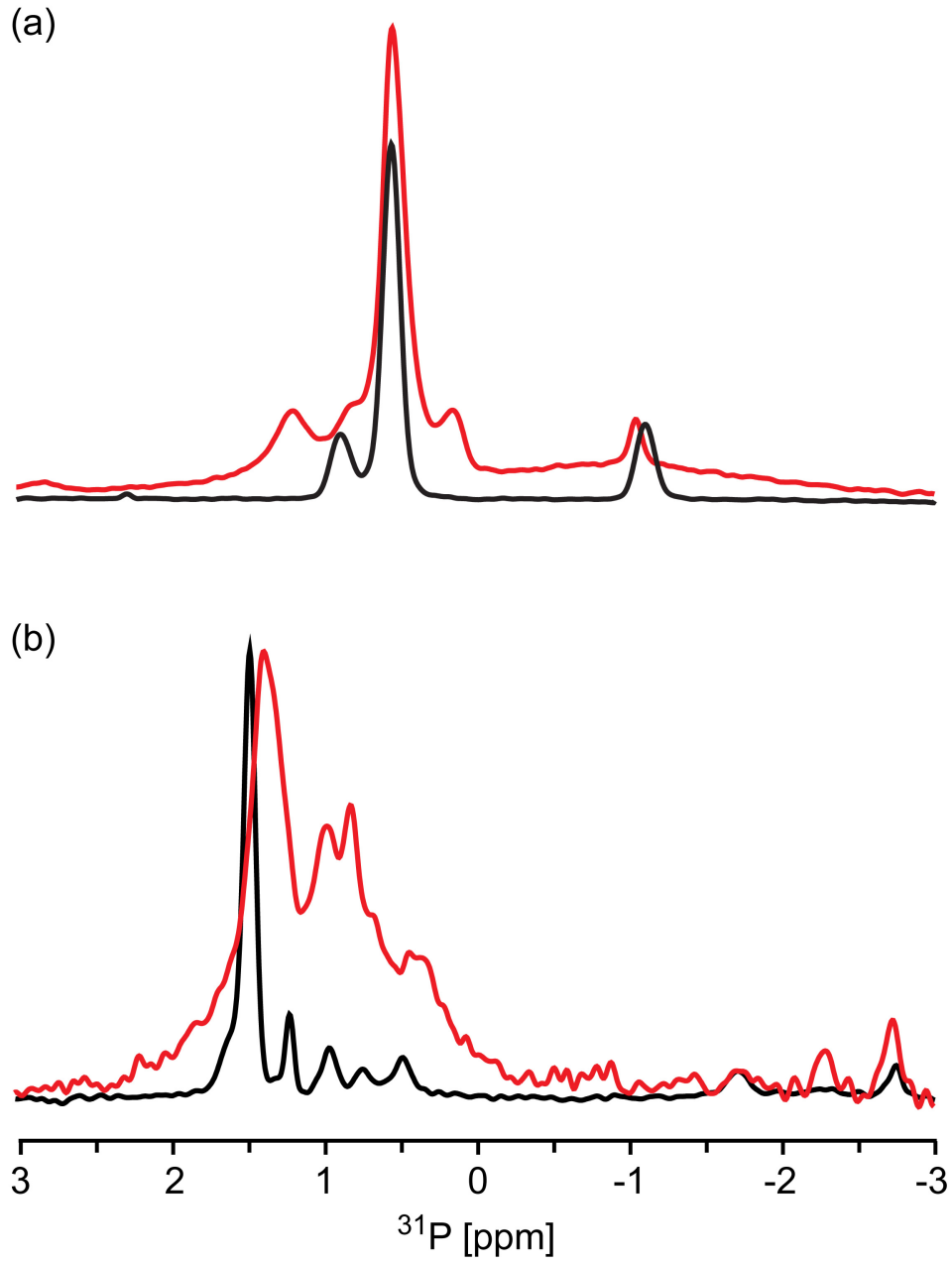


Figure 64 – Comparison the ^{31}P spectra recorded on intact cells (red) and isolated cell wall samples (black) from (a) *B. subtilis* 168 and (b) *B. subtilis* W23.

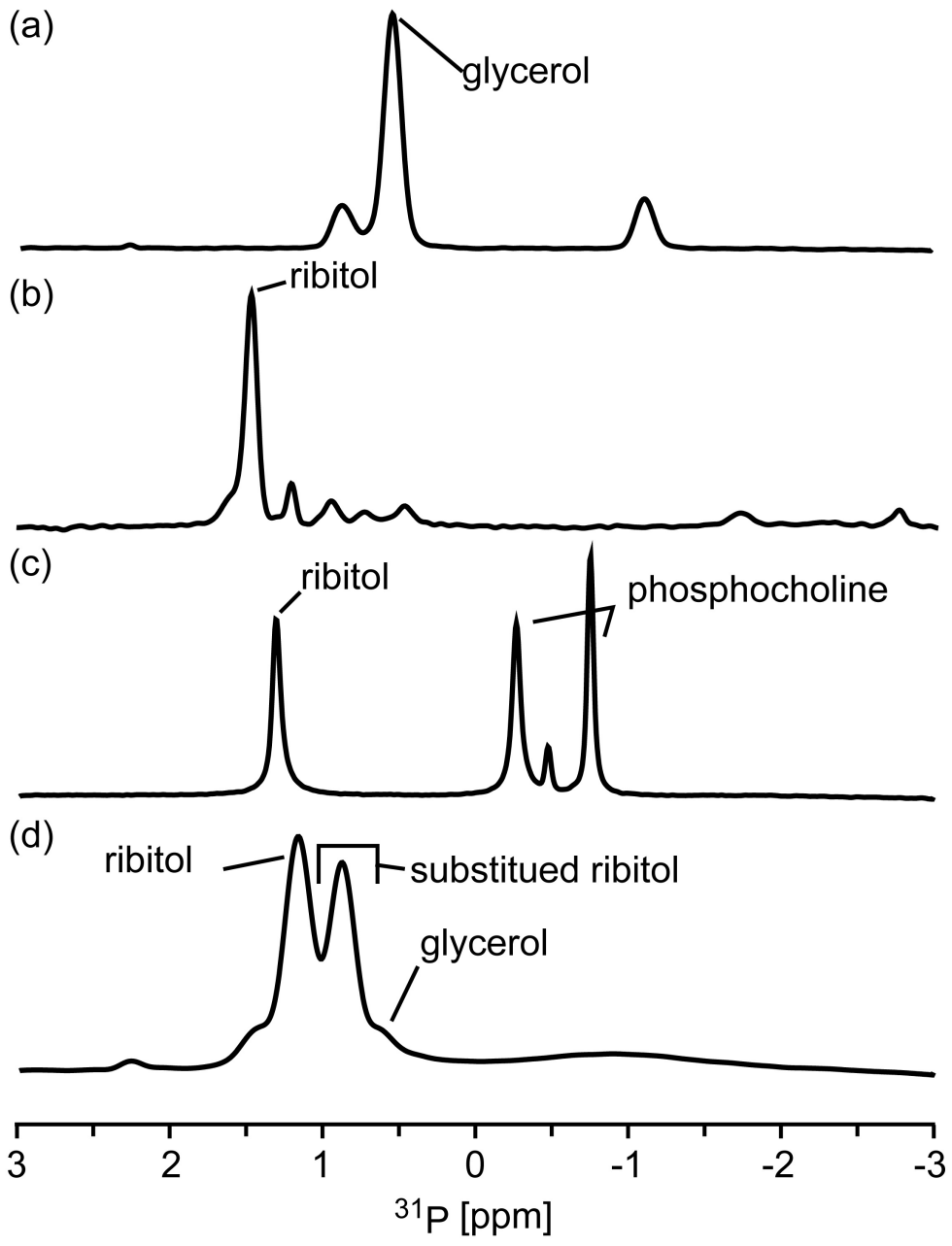


Figure 65 – Spectra collected on isolated cell wall samples of (a) *B. subtilis* 168, (b) *B. subtilis* W23, (c) *S. pneumoniae* R6 and (d) *S. aureus*. The assignment of phosphorous spectra is given in the figure.

1.4 to 0.6 ppm), whose WTAs would be predicted to contain ribitol- and glycerolphosphates substituted to variable degrees.

8.2.1 *Discussion*

The behavior of ^{13}C WTA resonances with respect to proton decoupling and CP dynamics suggests that a significant portion of the WTAs is restricted in its dynamics by the peptidoglycan network. Although accurate determination of the flexibility of WTA was not possible due to multiple overlaps in the 1D spectrum, this finding indicates that a part of the teichoic acid chains is localized outside the peptidoglycan network. This interpretation of the data is supported by cryo-EM images of whole cells or thin sections which suggest a "furry" surface on the cell that was supposed to be the wall- or lipid-linked TAs [110].

WTAs and LTAs as well as degraded fragments from various bacterial species have been previously investigated by ^{31}P liquid-state NMR spectroscopy [6, 77, 139, 150]. The only previous solid-state NMR study of TA analyzed the interaction of dried LTA with titanium surfaces at low MAS speed [137]. We show here that performing ^{31}P experiments on hydrated cell walls or intact bacterial samples at a reasonable MAS speed ($\approx$ 10 kHz) provides solid-state NMR spectra of a similar resolution to that found in liquid-state NMR. The advantage of the ^{31}P nucleus over the ^{13}C nucleus is its 100% natural abundance (vs 1.1% only for ^{13}C) as well as its appreciable sensitivity (4.2 times that of ^{13}C) making solid state ^{31}P NMR an easy and quick tool for further investigations of cell walls of Gram-positive species without the need of isotopic labeling. The comparative study on a variety of bacteria detailed in Figure 65 demonstrates that the rapid acquisition of a solid-state 1D ^{31}P NMR spectrum on hydrated intact bacteria (or processed cell walls with intact teichoic acids) allows easy identification of fine differences in the cell wall structure of bacterial species. The application of this method should provide a interesting tool to determine the effect of external factors on the WTA biosynthesis and chemical modifications, as well as having relevance in species identification and typification.

8.3 METAL BINDING TO WALL TEICHOIC ACIDS

One of the functions attributed to teichoic acids is a role in ions trafficking. To analyze the interaction of metal ions with WTAs in the presence of the

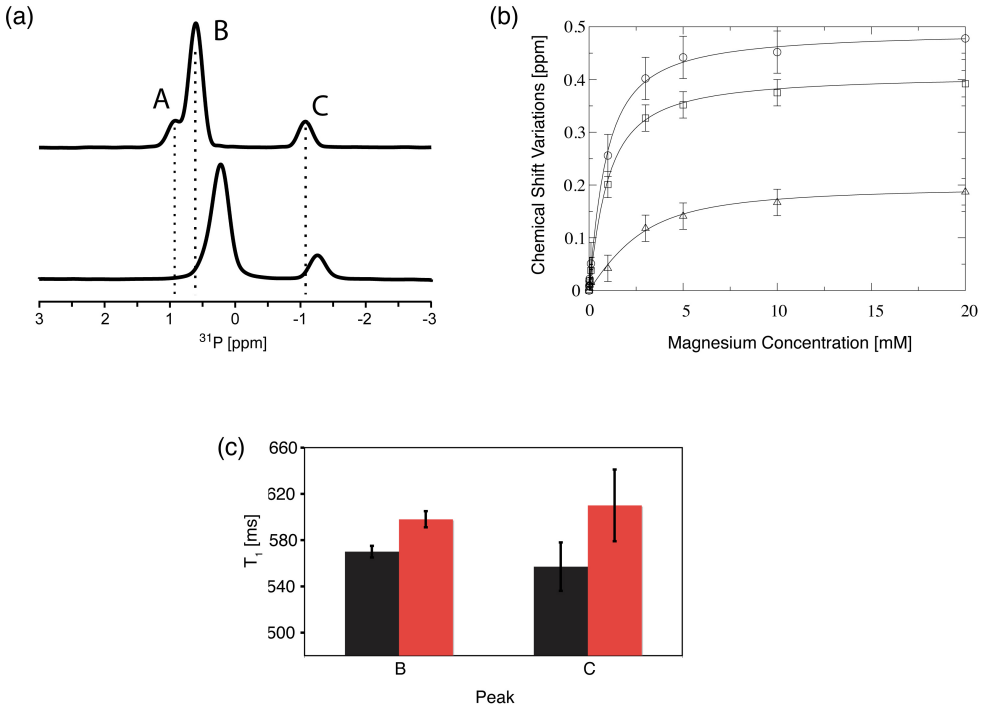


Figure 66 – Quantitative analysis of the interaction of magnesium ions with the phosphate groups of the cell wall WTA from *B. subtilis* 168. Panel (a) shows the 1D ^{31}P spectra in the absence of magnesium (top trace) and in the presence of 20 mM $MgCl_2$ (bottom trace). Panel (b) shows the ^{31}P chemical shift variations extracted from 1D ^{31}P spectra as a function of magnesium ion concentration for each individual ^{31}P resonance peaks A (open circle), B (square) and C (triangle). Curve fitting yielded Kd values of (595 $\pm$ 232) µM for peak A, (686 $\pm$ 321) µM for peak B and (1242 $\pm$ 730) µM for peak C. (c) ^{31}P longitudinal relaxation time constants (black : without Mg^{2+}, red : with Mg^{2+}) measured for each individual ^{31}P signal by inversion-recovery experiments collected on a Varian 600 MHz spectrometer with a 10 kHz spectral width, direct phosphorus excitation, 20 ms acquisition time and an interscan delay of 5 s. During acquisition TPPM ^{1}H decoupling was used at a proton rf-field strength of 50 kHz. The difference in T_1 values measured in absence/presence of Mg^{2+} is very small, but reproducible. For the most intense peak B, the difference in the time constants is larger than the estimated error in measuring T_1. For the less-intense peak C, the error is larger, but the tendency of longer T_1 for the cell wall samples in presence of Mg^{2+} could be reproduced.

peptidoglycan, we used the ^{31}P NMR probe. Figure 66a compares ^{31}P spectra recorded on a *B. subtilis* 168 cell wall sample first washed with EDTA (top trace) and then incubated with 20 mM $MgCl_2$ (bottom trace). This figure shows that the addition of Mg^{2+} ions produces chemical-shift variations of the WTA ^{31}P resonances. As the chemical shift variations appeared to be dependent on the concentration of $MgCl_2$, further samples of *B. subtilis* cell walls were titrated with increasing concentrations of Mg^{2+} (Fig. 66). The effect of Mg^{2+} interaction on WTA flexibility was investigated by measuring ^{31}P T_1 longitudinal relaxation time-constants in the absence and the presence of an excess (20 mM) of divalent ions (Fig. S1b). Higher T_1 relaxation rates indicate a lower flexibility as it has been shown by the measurement of relaxation time constants at different temperatures (see annex D). A very small increase of the relaxation time constant values is observed in the presence of Mg^{2+}, indicating a small decrease of the flexibility of WTA induced by the interaction with the magnesium ions.

Similar effects could also be observed for manganese ions. Indeed, ^{31}P spectra acquired on *B. subtilis* 168 cell wall samples, which were not treated with EDTA and grown in the presence of Mn^{2+} ions, showed significant line broadening of the ^{31}P WTA resonances compared to EDTA-treated samples. This broadening could only be explained by an interaction between the manganese paramagnetic ions and the WTA phosphate groups. Nevertheless it was impossible to obtain further information about the nature (electrostatic versus complexation) and location of the interaction site using ^{31}P NMR, although the more abundant glycerolphosphate groups seem also to be the more affected. To learn more about the interaction of these metal ions with the cell wall, ^{13}C correlation experiments were performed, as this nucleus has a good chemical-shift dispersion and its relaxation properties are sensitive to complexation with paramagnetic ions like Mn^{2+}.

Figure 67 shows the impact of the presence of residual Mn^{2+} ions on the ^{13}C-^{13}C through-bond correlation spectra of a freshly purified *B. subtillis* cell wall sample. Mn2+ ions were trapped by the cell wall during the bacterial culture from the growth medium containing 0.1 mM $MnCl_2$, and were retained through all washing and purification steps applied during sacculi preparation. As a result ^{13}C correlations of the glutamic acid (CO_α) and meso-A_2pm carboxyl groups were significantly broadened such that they essentially disappeared from the correlation spectrum (Fig. 67). Interestingly, the interaction between the peptidoglycan and manganese ions only occurs significantly in the presence of WTAs (Fig. 67). The concomitant resonance broadening observed in ^{13}C and ^{31}P spectra thus suggests a close proximity

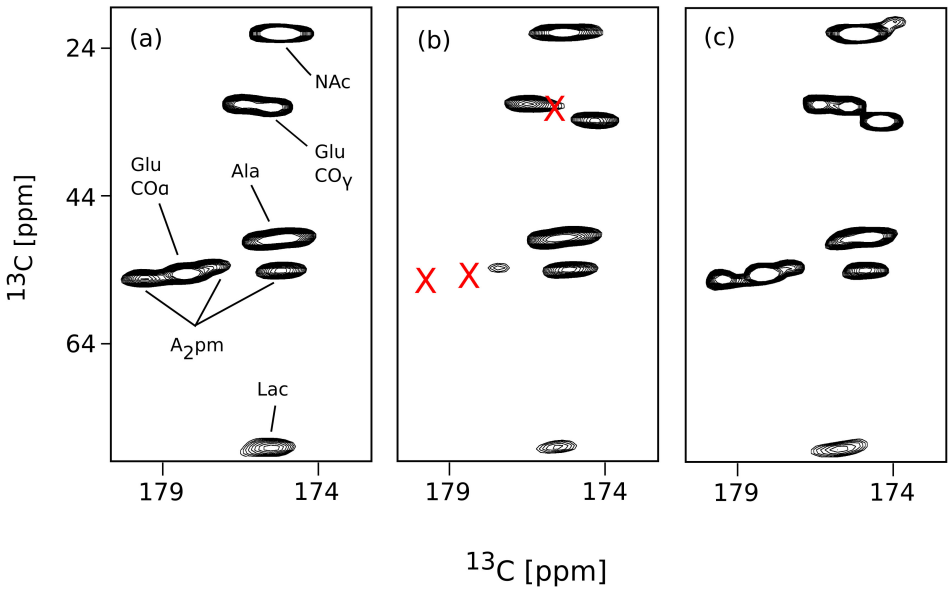


Figure 67 – Mn^{2+}-induced broadening effects observed on the peptidoglycan and WTA resonances of *B. subtilis* 168 cell-wall using ^{13}C-^{13}C correlation experiments collected on a 400 MHz Bruker spectrometer. Figure shows the correlations between carbonyl groups and aliphatic carbons of peptidoglycan samples prepared (a) without WTA and without EDTA treatment, (b) with WTA and without EDTA treatment and (c) with WTA and after EDTA treatment. Carbonyl correlations of the Glu and meso-A_2pm residues are broadened by the interaction with Mn^{2+} ions only in the presence of WTA.

of the paramagnetic Mn^{2+} ions to the glutamic acid and to the meso-A_2pm carboxyl groups of the peptidoglycan on one side, and to the phosphate groups of the WTAs on the other side.

8.3.1 *Discussion*

Interaction between anionic TAs (or negatively charged TAs) and cationic divalent ions has been demonstrated and largely discussed in the past [6, 88, 89, 140]. Data obtained by dialysis exchange or liquid-state NMR on isolated LTAs or extracted WTAs indicated a relatively low affinity for Mg^{2+} (Kd $\approx$ 15 mM) as well as for Ca^{2+} (Kd $\approx$ 50 mM). A monodentate coordination of the magnesium ion by the teichoic acid phosphates has been proposed based on the analysis of magnesium to phosphorus 1:1 ratio. The affinity of teichoic acids for Mg^{2+} is comparable to the one for free HPO_4^{2-} ions. However, a lower dissociation constant of approximately 350

μ M has been measured on intact *S. aureus* cell wall and the possibility of a bidentate coordination mode for Mg^{2+} by one or two different phosphate groups of the poly(glycerolphosphate) teichoic acid chain has been proposed [88, 140, 182]. Solid-state NMR allows the study of metal ion interactions within the complete cell wall (containing WTAs and peptidoglycan) and the role of the cell wall as coordination sphere of the divalent ions. The analysis of the ^{31}P chemical-shift variations as a function of the magnesium concentration leads to a dissociation constant of about 600 ± 300 μ M for the phosphate of the repeating unit (Fig. 66). Thus Mg^{2+} exhibits a stronger interaction with intact cell wall than with isolated TAs (Kd ≈ 15 mM) suggesting that the architectural arrangement of the WTA polymer in the cell wall could form an integrated system for supplying cations to the membrane. Remarkably, the presence of Mn^{2+} ions in cell wall samples was detected through line broadening of the corresponding ^{31}P and ^{13}C NMR resonances which is expected for paramagnetic atoms. Mn^{2+} ions were present in the growth medium at a concentration of 0.1 mM, and they were not removed from the cell wall by several washing steps with HEPES pH 7.5 buffer. Only a treatment with EDTA, which forms complexes with divalent ions with very low dissociation constants (1.62 nM for Mg^{2+} and 13.49 fM for Mn^{2+}), was able to remove the metal ion from the cell wall and to allow the recovery of a sharp ^{31}P line width. From the large line-broadening of the ^{31}P resonances observed at very low Mn^{2+} concentrations, the dissociation constant of this ion with WTAs can be postulated to be as low as, if not even lower than, the binding constant determined above for Mg^{2+}.

Interestingly, ^{13}C correlation spectra of the cell wall of *B. subtilis* taken directly after purification from a growth medium containing 0.1 mM Mn2+ ions without EDTA treatment showed the loss of the glutamic acid and meso-A_2pm carboxyl group resonances. Analyzing the ^{13}C and 31P data together and taking into account that binding of Mn^{2+} ions does not occur significantly in the absence of WTAs, we propose an intermolecular complexation of the Mn^{2+} ion by the carboxyl groups of glutamic acid and meso-A_2pm of the peptidoglycan and by the phosphate groups of the teichoic acids (Fig. 68). The Mn^{2+} metal center is usually found in an octahedral geometry with oxygen donor ligands. The peptidoglycan carboxylate groups from γ-D-iso-Glu and meso-A_2pm can provide a bis(η^2-carboxylato) scaffold for the manganese complexation, which can be nicely complemented with a mixture of monodentate phosphate groups of the teichoic acids repetition units and of water molecules or alternatively by a bidentate phosphate ligand [88, 182]. Taken together these results provide the first experimental

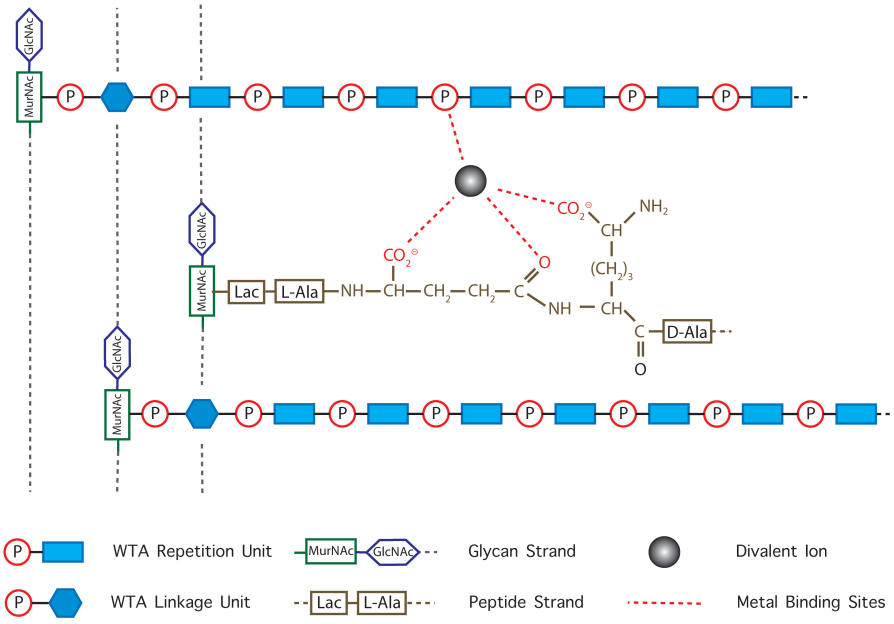


Figure 68 – Schematic representation of possible binding sites for divalent metal ions on the example of the cell wall of *B. subtilis*. Red dotted lines show cooperative metal binding between the peptide strands and the phosphate groups of the WTA. Lines from the carboxylate or the phosphate groups could represent either monodentate or bidentate ligation modes.

evidence for WTA and peptidoglycan acting in a cooperative role in the geometry and the thermodynamics of the binding to divalent cations. It also gives an explanation for the higher association constant of Mg^{2+} in the cell wall when compared with isolated TAs. Amidation of carboxyl groups prevents interaction with metal ions. In *B. subtilis*, the amidation of the carboxyl group of A_2pm is modulated as a function of the growth conditions and thus could represent a mechanism to regulate the amount of cell wall bound ions. For *S. aureus*, where the glutamic acid is always amidated, Mg^{2+} binding might be controlled by the rate of alanyl substitution in the TAs as discussed in the literature [88, 124]. Inverse correlation between the rate of alanylation and Mg^{2+} concentration has been determined for different growth conditions [58, 62]. The proposed peptidoglycan-WTA mediated binding of cations could be a key factor for storage and supply of cations essential for bacterial growth [151]. It has also been assumed that the isoelectric point of the cell wall determines the susceptibility of the cell to cationic antimicrobial host defense molecules, bacteriocins and antibiotics [136] and could play a role in cell adhesion [138]. A modification of the compactness of the peptidoglycan has been described in the case of alanylation of the TAs

due to modification of the charge distribution. Cation-mediated interactions between peptidoglycan and WTA, as described here, could represent an important structural aspect of Gram-positive cell wall architecture. It has been proposed that for Gram-positive bacteria, the absence of an external membrane is compensated by a thicker peptidoglycan whose porosity is reduced by the TAs [180].

8.4 CONCLUSION

^{31}P solid-state NMR emerges as a powerful tool to study the wall teichoic acids in isolated cell walls as well as in intact cell samples, as it does not require any specific labeling. It is particularly sensitive to the presence of divalent metal ions and affords a way to identify the binding sites as well as to efficiently determine the binding constant. ^{31}P spectroscopy is nicely complemented with ^{13}C NMR on uniformly labeled samples that yield information on the attached peptidoglycan as well as on the saccharide and alditol content of the teichoic acids. The combination of the two approaches allowed proposing a new geometry for the coordination of different divalent metal ions.

9

PERSPECTIVES

Peptidoglycan structure and polymorphism

Since protein recognition motifs are not only defined by the chemical constitution of peptidoglycan subunits but also by their three-dimensional organization in the polymer, the absence of structure at atomic resolution constitutes a stumbling block to a better understanding of the molecular recognition mechanisms and of the catalytic pathways used by peptidoglycan-interacting proteins. Therefore our main goals is the determination of three-dimensional peptidoglycan structure. As shown in the chapter 5, conflicting models of the macroscopic structure of peptidoglycan have been built in silico based on the known structure of the subunit and biophysical properties of the polymer. Strikingly, the debated issues include basic features such as the orientation of the glycan strands that could be either parallel or perpendicular to the cell membrane. The mode of insertion of novel subunit in the expanding peptidoglycan network also remains highly speculative. In this context, solid-state NMR could build the first experimental model of bacterial peptidoglycan 3D-structure.

The mode of peptidoglycan synthesis of is globally conserved in bacteria, although important polymorphisms have been detected including both interspecies variations and acquisition of resistance to antibiotics. These polymorphisms include the thickness of the peptidoglycan and various modes of peptidoglycan cross-linking. In chapter 8 and 9 it was demonstrated that solid-state NMR is suitable to explore the impact of these polymorphisms on the global architecture of peptidoglycan. The next objects will be the unusual reticulation of the *Micrococcus luteus* peptidoglycan, and substitution of 4-3 by 3-3 cross-links in *Enterococcus faecium* mutants resistant to β-lactam antibiotics.

Teichoic acids structure and flexibility

Further, knock out mutants of proteins of the TAG familly involved in the biosynthesis of the teichoic acids have already been prepared, and our methods will be used to determine changes in the chemistry and flexibility of the peptidoglycan and teichoic acids due to the modification on the teichoic acids. It is also known that cells from *B. subtilis* mutants lacking teichoic acids due to TagO knockouts can be viable under certain conditions. In this context we will study the influence of missing teichoic acids on the peptidoglycan. All TAG proteins are generally interesting targets for the development of new antibiotics as the teichoic acids are an important structure for bacterial survival.

Cell wall recognition by proteins

Another of our perspectives is to go further in the interaction studies involving large fragments or the entire peptidoglycan/cell wall polymer to analyze protein recognition of the cell wall or peptidoglycan. In this context we are interested in several interaction systems.

SYNTHESIS AND DEGRADATION OF THE PEPTIDOGLYCAN: The ongoing emergence of resistance bacterial strains has brought the activity of the transptidases, main targets of the antibiotics into the center of interest. In the last year we have setup protocols for the expression of isotopically labeled forms of two transpeptidases, a D,D-transpeptidase and a L,D-transpeptidase. Both proteins have been completely assigned and interaction studies with the peptidoglycan of *B. subtilis* are in progress. Structures of both proteins are already known, and the focus will be on refining structural modification of the proteins upon binding with fragments of the peptidoglycan from different strains and with antibiotic. As L,D-transpeptidases are generally not inhibited by β-lactam antibiotics, finding active antibiotics against this class of proteins would allow to inhibit a potential escape mechanism of resistant bacteria [103, 104]. Until now no interaction with fragments of the peptidoglycan have been observed, due certainly to the weak affinity of the transpeptidases to muropeptides. The concepts we have used allow to observe protein binding directly on the entire cell wall and have potentially broad applications beyond the study of autolysis and peptidoglycan morphogenesis. They are also applicable to the study of eukaryotic peptidoglycan recognizing proteins involved in the innate immunity (PGRPs).

IN BACTERIA NMR: Modulation of the interactions for different physiological conditions can only be assessed by recording data on a complete bacteria sample. This motivates the extension to the study of intact sacculi in a physiological environment. In addition to recognition phenomena at the molecular level, morphogenesis and cell-wall development along the cell cycle could be studied, allowing the description of cellular events at the molecular and even atomic level. We have already acquired ^{31}P and ^{13}C spectra of complete bacteria samples (as shown for the ^{31}P spectra in chapter 9), and the possibilities to study molecular interaction in intact bacteria open a large number of potential applications. These methods can also contribute to the study of the interaction of bacteria with host cells as many virulence factors are bound to peptidoglycan.

Part VI

ANNEXES

A

MATERIAL AND METHODS

A.1 PEPTIDOGLYCAN PURIFCATION

A.1.1 *Cell cultures in isotopically labeled media*

Bacillus subtilis subsp. subtilis strain 168 cells from the American Collection (ATCC 23857) and *E. coli BL21 (DE3)* were grown in standard ^{13}C,^{15}N-labeled M9 medium containing 0.1 mM $MnCl_2$ (see annex). *Staphylococcus aureus* (ATCC 25923) cells were grown in a ^{13}C,^{15}N-SILANTES medium (www.silantes.com). *Streptococcus pneumoniae* R6 was grown in C+Y medium [60]. *Bacillus subtilis* subsp. *spizizenii* W23 cells were grown in standard unlabeled LB medium. For the cell cultures in M9 medium a supplementary preculture step is necessary to acclimatize the cells to the poor medium. Cells were harvested by centrifugation at an OD (600 nm) of about 0.6-0.7 at 4◦C to avoid degradation. Highly purified peptidoglycan with and without teichoic acids can be obtained by previously published procedures [43, 157].

A.1.2 *Isolation of the peptidoglycan from isotopicaly labeled E. coli cell cultures*

Volumes given are reference to 1l of cell culture harvested at an OD_{600} of about 0.6. The cells are retaken in 20 ml ice-cold miliQ water and dropped in 20 ml of a boiling 8% w/v SDS solution. After 30 min of boiling the cell extract is kept at room temperature until its cooled down and the pellet is recovered by ultracentrifugation (30 min, 46000 g).

The sample is washed SDS free be repeated resuspending in 40 ml of tepid miliQ water followed by ultracentrifugation (30 min, 46000 g) until the Hayashi control [55] is negative.

The final SDS free pellet is retaken in 6.75 ml of a 10 mM Tris/HCl, 10 mM NaCl solution at pH 7. 750 µl of 3.2 M imidazol at pH 7 is added to inhibit potential lysozyme activity.

10 mg/ml α-Amylase is added and the sample is incubated at 37°C to cleave glycogens from the cell wall. After 2 hours 10 mg/ml of Pronase E is added to remove lipoproteins. After one additional hour of incubation at 60°C SDS is added to a concentration of 4% w/v and the sample is incubated at 80°C for 30 min to stop the enzyme's activity.

Once the suspension is cooled down to room temperature the SDS is again removed by washing steps in miliQ water. The SDS free pellet is then retaken in 0.02% w/v NaN_3 and can be stored at 4°C.

A.1.3 *Isolation of the cell wall from isotopicaly labeled cell cultures of Gram-positive bacteria*

Harvested cells are retaken in 30 ml of an ice-cold 50 mM Tris/HCl buffer at pH 7 and boiled in 120 ml of a 5% SDS solution for 30 min. After boiling they are washed twice in a first step with 20 ml of a 1M NaCl solution by resuspension and ultracentrifugation, before they are washed SDS free with miliQ water as described for *E. coli* in the previous subsection. The SDS free sample is then retaken in 6-10 ml miliQ water and 12-16 20 sec Fastprep steps are applied. The sample is softly centrifuged for 5 min and the supernatant is retaken for ultracentrifugation (30 min, 46000 g).

The sample is then retaken in 10 ml of a 100 mM Tris/HCl, 20 mM $MgSO_4$ buffer at pH 7.5. 10 µg/ml of DNAse and 50 µg/ml of RNAse are added and the sample is incubated at 37◦C. After 2h 10 mM of $CaCl_2$ and 100 µg/ml of Trypsin are added for protein digestion. The sample is kept in incubation for further 18 h.

After incubation the sample concentration is brought to 1% w/v of SDS to inactivate the enzymes and incubated for 15 min at 80°C and the pellet is recovered by ultracentrifugation (30 min, 46000 g).

The pellet is then retaken in 10 ml of 8 M LiCl and incubated for 15 min at 37°C. One washing step in miliQ water is applied followed by a washing step in acetone, before the sample is finally washed SDS free and stored at -80°C prior to freeze drying.

A.1.4 *Teichoic acids removal from purified cell walls*

5 mg of the freeze dried cell wall sample is retaken in 3-5 ml concentrated HF and incubated for 48 h at 4°C. The sample is then ultracentrifugated (30 min, 46000 g) and the pellet is washed twice in 20 ml ice-cold miliQ water, followed by two washing steps in 20 ml of 100 mM Tris/HCl at pH 7. Finally the pellet is retaken in 0.02% NaN_3 and can be stored at 4◦C

A.1.5 *Muropeptide purification from isolated cell walls*

500 µl of peptidoglycan (murein) suspension (≈5 mg/ml) is complemented with 200 µl 80mM NaP (pH 4.8), 40 µl 260 µ/ml Cellosyl, and 60 µl miliQ for digestion at 37◦C during 24h, if needed 40 µl of Cellosyl is added after 24 h for further digestion. The reaction is stopped by boiling at 100°C for 10 min. The sample is stored at 4°C.

A.2 ROTOR PREPARATION FOR SOLID-STATE NMR

PEPTIDOGLYCAN: The purified peptidoglycan was spun down by centrifugation in a 5415D Eppendorf centrifuge at 18000 rpm for 15 min at 25¡C and resuspended in a 50 mM HEPES buffer at pH 7.5. These steps were repeated until the pH of the supernatant remained stable at 7.5. A hydrated gel-like suspension of peptidoglycan was packed in the rotor by centrifugation.

CELL WALL : Lyophilized cell walls were resuspended in HEPES buffer pH 7.5 and were treated in the same manner than the peptidoglycan. Complete removal of metal ions for reference peptidoglycan and cell wall samples was achieved by a 500 mM EDTA treatment before conditioning of the samples in HEPES (metal free) buffer.

MG^{2+} TITRATION EXPERIMENTS AND LONGITUDINAL RELAXATION MEASUREMENTS IN PRESENCE AND ABSENCE OF MAGNESIUM IONS: A metal ion free cell wall sample of *B. subtilis* 168 (ATCC 23857) was separated into

10 equivalent fractions which where then spun down. The pellets were resuspended in 1 ml of 50 mM HEPES buffer pH 7.5 containing different concentrations of Mg^{2+} (0, 10 µM, 20 µM, 30 µM, 100 µM, 1 mM, 3 mM, 5 mM, 10 mM, 20 mM) and 1% HMPA for spectra referencing. The samples at 0 and 20 mM Mg^{2+} concentrations were also used for ^{31}P longitudinal relaxation measurements.

A.3 SOLID-STATE NMR EXPERIMENTS

NMR experiments were performed at 9.4 T (^{1}H NMR frequency at 400.13 MHz) using a Bruker 400 MHz spectrometer equipped with a Bruker 4 mm triple-resonance probe and at 14.1 T (^{1}H NMR frequency at 600 MHz) using a Varian spectrometer equipped with a 3.2 mm triple resonance BioMAS probe.

A.3.1 ^{31}P *1D experiments*

Spectra were obtained by direct ^{31}P excitation. For the spectra acquired at 400 MHz the phosphorus rf-field strength was set to 60 kHz. During acquisition SPINAL or CW heteronuclear decoupling was used at a proton rf-field strength of 90 or 45 kHz, respectively. Acquisition time was between 50 and 120 ms, depending on the decoupling field strength. The experimental time was 1 h with an interscan delay of 5 s. A cosine square apodization function was applied prior to Fourier transformation. For the Mg^{2+} binding experiments done on the 600 MHz Varian spectrometer the spectral width was set to 10 kHz. The acquisition time was 20 ms and the interscan delay 1 s. Experimental time was about 15 min per titration point. To determine the affinity of Mg^{2+} for *B. subtilis* cell walls from the experimental data, the ^{31}P chemical-shift variations were measured as a function of the Mg^{2+} concentration. Curves of chemical-shift variations as a function of Mg^{2+} concentration have been obtained for the different ^{31}P resonances and plotted. Dissociation constants were extracted by fitting the curves independently using the classical 1:1 binding equation [12].

^{31}P *longitudinal relaxation time-constants*

Longitudinal relaxation time-constants were measured by inversion recovery after direct excitation. The interscan delay in the later experiments was increased to 5 s. A cosine square apodization function was applied prior to

Fourier transformation. During acquisition TPPM heteronuclear decoupling was used at a proton rf-field strength of 100 kHz and the phosphorus rf-field strength was set to 50 kHz.

A.3.2 ^{31}C *experiments*

^{13}C *through-bond correlation experiments*

The pulse sequence used for through-bond ^{13}C correlation experiment on the peptidoglycan of *E. coli* was based on the solid-state refocused INADEQUATE sequence [92], whose phase cycle was modified to avoid the signal reduction of a factor two due to the double-quantum selection. The result is very similar to the recently published UC2QF COSY experiment [16]. The ^{1}H-^{13}C cross-polarization (CP) behavior being very unlike for different resonance in the peptidoglycan, we preferred to start the experiment with direct excited ^{13}C magnetization instead of using CP. This introduces a slightly loss of sensitivity for aliphatic carbons when the same repetition delay is used. However, the sensitivity of CO resonances is greatly improved and the correlation spectra do not show the intrinsic asymmetry observed in multidimensional solid-state NMR correlation spectra involving CP [14]. Moreover, the peak-intensity comparison between correlation spectra of peptidoglycan alone or in interaction with the protein are not biased by potentially different CP dynamics.

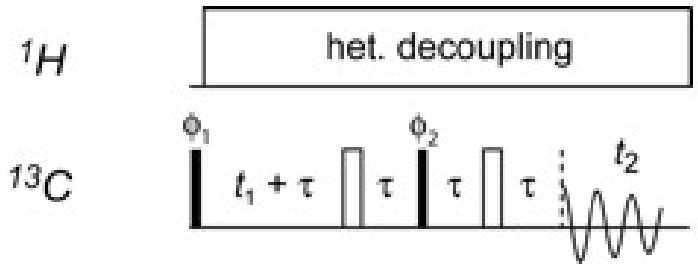


Figure 69 – Pulse sequence for through-bond ^{13}C correlation experiment. 90° and 180° rf pulses are represented by filled and open bars, respectively. All pulses are applied along the x axis unless indicated. A four-step phase cycle was applied with ϕ_1 = x, -x; ϕ_2 = (2y), 2(-y); ϕ_{rec} = x, -x, x, -x.

All ^{13}C through-bond correlation experiments on *E. coli*, *B. subtilis*, and *S. aureus* presented were performed on the 400 MHz Bruker spectrometer. The spectrum presented for *S. pneumoniae* was recorded on the 600MHz Varian spectrometer.

EXPERIMENTS RECORDED AT 400MHZ The real sample temperature was kept at about 7 °C, and the MAS frequency set at 9 kHz. Initial ^{13}C magnetization was obtained by direct carbon excitation. The ^{13}C through-bond correlation experiments performed on samples of *B. subtilis* and *S. aureus* were recorded as for *E. coli* except that initial ^{13}C magnetization was obtained by an adiabatic cross-polarization step of 500 ms [57], performed at an averaged rf-field strength of 69 kHz on 1H and 60 kHz on ^{13}C (also used for ^{13}C pulses). Both J-evolution periods were rotor synchronized. The delay τ was adjusted in an 1D-INADEQUATE experiment to give maximum transfer efficiency and was set to 1,6666 ms for *E. coli* and to 2 ms for *B. subtilis* and *S. aureus*. SPINAL ^{1}H decoupling [37] was used with ω_1 = 89 kHz during J-evolution and acquisition. The spectral width was 24 kHz in both dimensions. Acquisition times were 15-18 ms in the direct dimension, and 5-7 ms in the indirect dimension using States-TPPI quadrature detection. The interscan delay was set to 1.5 s. for *E. coli* and *S. aureus*, or 2 s for *B. subtilis*, leading to a total experimental time between 18 and 62 h depending on the bacterial strain. A cosine square apodization function was applied in both dimensions prior to Fourier transformation.

EXPERIMENTS RECORDED AT 600MHZ The sample temperature was kept at about 12°C, and the MAS frequency set at 12.5 kHz. As for *B. subtilis* and *S. aureus* ^{13}C magnetization was obtained by direct excitation, the transfer delay τ was set to 1.666 and TPPM decoupling at ω_1 = 66kHz was used during J-evolution and acquisition. Spectral width used were 40kHz for the direct dimension and 30kHz for the indirect dimension. The acquisition time was set to 10 ms in both dimensions direct and a interscan delay of 1.5 s was used leading to a final experimental time of 63h. A cosine square apodization function was applied in both dimensions prior to Fourier transformation.

^{13}C *through-space correlation experiments*

Initial ^{13}C magnetization was obtained by adiabatic cross polarization. The indirect acquisition time t_1 was set to 9.5 ms, 10 ms, and 6 ms for (a), (b), and (c), respectively. Three different mixing times τ_{mix} = 10 ms, 30 ms, and 100 ms were used to analyze the flexibility of the peptidoglycan. States quadrature dectection was applied in the indirect detection dimension. Other experimental details are the same as for the through-bond correlation experiment (see above).

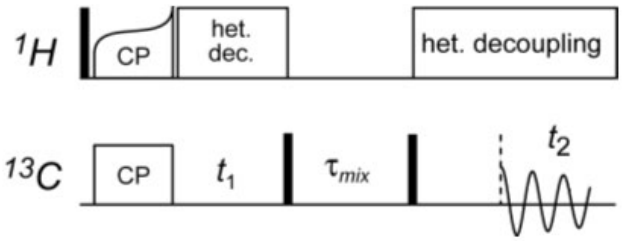


Figure 70 – PDSD sequence for through-space ^{13}C correlation experiment.

^{13}C longitudinal relaxation time-constants

Carbon longitudinal relaxation were measured by inversion-recovery after direct excitation on the 400MHz Bruker spectrometer. The interscan delay has been raised to 5 s to assure optimal polarization prior to the following scan. Other parameters were set as described above for the experiments run under 400MHz.

A.4 PROTEIN-PEPTIDOGLYCAN INTERACTION

For the peptidoglycan reference spectrum the rotor was prepared as previously described (see section "Rotor preparation for solid-state NMR"). To test the sample preparation for the study of protein interaction, four samples of a total volume of 100 µl containing a YajG concentration of 20µM and different volumes of peptidoglycan preparation (68µl, 50µl, 25µl, 12.5µl) were incubated four 2h at 37°C. The samples were centrifuged in a Eppendorf tube at 46000g and 37°C for 30min. 5µl of each samples supernatant were loaded on the SDS-Page gel (Figure 71). The precipitate was retaken in 100µl Laemmli loading buffer (4x) and incubated at 90°C for 5min. It was then centrifuged in 5415D Eppendorf centrifuge at 14000rpm for 10min. 5µl of each samples supernatant were loaded on the SDS-Page gel. The gel shows that protein concentration in the precipitate is rising with the concentration of sacculi.

For the acquisition of the peptidoglycan-protein interaction spectrum, 50µl of the spun down cell wall preparation (5415D Eppendorf centrifuge, 18000rpm, 5min) were incubated in 1ml of a 800µM YajG protein solution (50mM MES, pH6) for 2h at 37°C. After incubation the sample was centrifuged (5415D Eppendorf, 18000rpm, 15min) and the precipitate (37mg) charged into the solid state. The sample was stored at 4◦C.

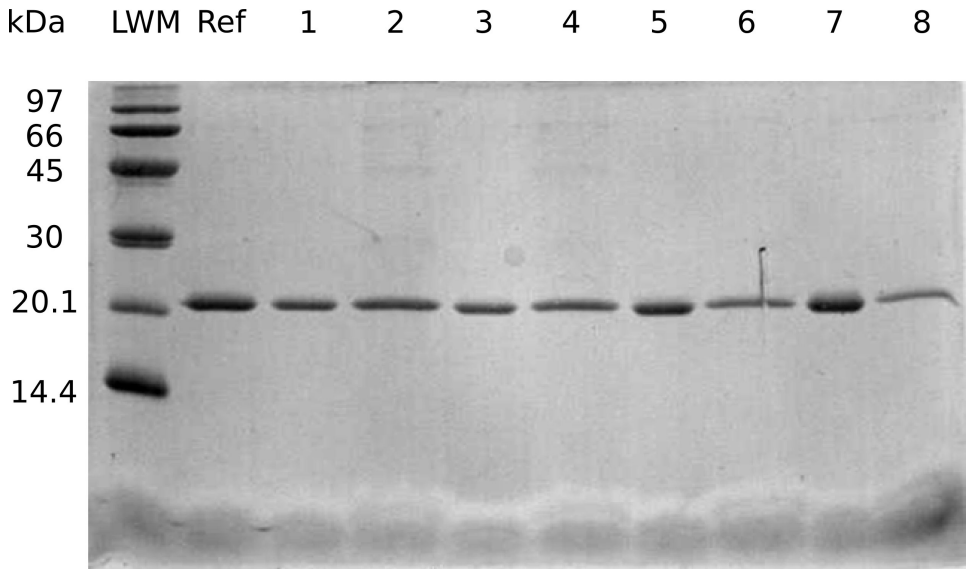


Figure 71 – SDS-Page control of YajG-peptidoglycan sample preparation. (1) Supernatant of the 68µl peptidoglcycan sample (2) Precipitate of the 60µl peptidoglycan sample (3) Supernatant of the 50µl peptidoglcycan sample (4) Precipitate of the 50µl (5) Supernatant of the 25µl peptidoglcycan sample (6) Precipitate of the 25µl (7) Supernatant of the 12.5µl peptidoglcycan sample (8) Precipitate of the 12.5µl

B

MINIMAL MEDIUM FOR ^{13}C, ^{15}N LABELLED CELL CULTURES ($M9^+$)

Table 3 – $M9^+$ minimal medium components

Base solution	Component	Concentration [mM]
	$Na_2HPO_4.7H_2O$	37.3 (10 g/l)
	KH_2PO_4	22 (3 g/l)
	NaCl	8.5 (0.5 g/l)
	$^{15}NH_4Cl$	18.6 (1 g/l)
Supplements	**Component**	**Concentration [mM]**
	^{13}C Glucose	22 (4 g/l)
	$MgSO_4$	1 (0.12 g/l)
	$CaCl_2$	0.1 (11 mg/l)
	$MnCl_2$	0.1 (12.5 mg/l)
	$ZnSO_4$	0.05 (mg/l)
Vitamines	**Component**	**Concentration [µM]**
	Pyridoxin	295 (50 µg/l)
	Biotin	205 (50 µg/l)
	Panthothenate, hemi calcium	228 (50 µg/l)
	Folic acid	113 (50 µg/l)
	Choline chloride	358 (50 µg/l)
	Niacin amide	409 (50 µg/l)
	Riboflavin	133 (5 µg/l)
	Thiamine	741 (250 µg/l)

C

PEPTIDOGLYCAN AND WALL TEICHOIC ACIDS ASSIGNMENT

C.1 ^{13}C peptidoglycan chemical shifts (in ppm) from different bacterial species

Table 4 – ^{13}C peptidoglycan chemical shifts (in ppm) from different bacterial species at 283K, pH 7.5 in 50mM HEPES.

		E. coli	B. subtilis 168	S. aureus	S. pneumoniae
GlcNAc					
	C1	100.9	101.0		
	C2	55.6	56.2		
	C3	72.3	72.8		
	C4	79.8	80.2		
	C5	75.2	75.3		
	C6	60.3	60.9		
	NHAc CO	174.9	175.2	175.4	
	NHAc CH_3	23.0	23.0	23.0	
MurNAc					
	C1	101.8	102.1		
	C2	54.8	54.9		
	C3	79.2	79.9		
	C4	72.2			
	C5	75.2			
	C6	60.3			

Table 4 – ^{13}C peptidoglycan chemical shifts (in ppm) from different bacterial species at 283K, pH 7.5 in 50mM HEPES.

	NHAc CO	174.9	175.2	176.0	
	NHAc CH_3	23.0	23.0	23.0	
	Lac-CO	175.0	175.2		
	Lac-C_α	78.1	78.7		
	Lac-CH_3	18.1	18.5		
L-Ala D-Ala					
	C_α	50.1	50.7	50.4	50.8/50.1
	C_β	17.5	17.3	17.1	17.5/17.9
	CO	174.9	175.9	176.0	176.0/173.0
γ-D-Glu					
	C_α	54.7	55.0	53.4	
	C_β	28.6	29.0		
	C_γ	32.3	32.5	32.0	
	CO_δ	175.8	176.4/175.5	175.5	
	CO^{2-} $CONH_2$	178.0	178.3		
meso-A_2pm or Lys					
	C1/C1′ or C_α	54.8/54.5	55.7/53.7	54.7	
	C2/C2′ or C_β	31.2	31.8/31.1	31.1	
	C3 or C_γ	21.5	22.5/21.5	22.9	
	C_δ				28.5
	C_ϵ			39.6	
	CO/CO′	174.4/173.6	179.5/177.5 175.5/172.9	175.1	

Table 4 – ^{13}C peptidoglycan chemical shifts (in ppm) from different bacterial species at 283K, pH 7.5 in 50mM HEPES.

Gly					
	C_α			172.5/171.4	
	CO			43.2	

C.2 ^{13}C TEICHOIC ACIDS CHEMICAL SHIFTS (IN PPM) FROM DIFFERENT BACTERIAL SPECIES

Table 5 – ^{13}C teichoic acids chemical shifts (in ppm) from different bacterial species at 283K, pH 7.5 in 50mM HEPES.

		B. subtilis ATCC 168[3]	*S. aureus*[1]	*S. pneumoniae*[2]
Ribitol[4]				
	C1		67.1 (66.8)	
	C2		71.2 (70.2)	
	C3		71.3 (71.7)	
	C4		80 (70.2)	71.5
	C5		65.3 (66.8)	67.4
Glycerol-P[4]				
	C1	66.1 (67.3)		
	C2	76.2 (71.6)		
	C3	65.4 (63.1)		
β-GlcNAc				
	C1		101.9	
	C2		56.2	
	C3		74.6	
	C4		70.6	
	C5		76.2	
	C6		61.2	
	NHAc CO		175.4	
	NHAc CH_3		23	

1 The major component of *S. aureus* teichoic acids is identified as a 1,5-poly-[4-*O*-(2-acetamido-2-deoxy-beta-glucopyranosyl)ribitol-diphosphate]. The minor component (chemical shifts within parentheses) contains an unsubstituted poly-phosphoribitol.

2 Ribitol spin system could not be completed due to major overlaps in the 69-72 ppm window. The stereochemistry at the C1 carbon of glucopyranoside is derived from its chemical shift value. The assignment of GlcNAc versus Glc is based on the C2 chemical shift value. Assignment of AAT-Gal is completely unambiguous with its unique C6 chemical shift value.

3 The major component of *B. subtilis* teichoic acids is identified as 1,3-poly-[2-(α-glycopyranosyl)glycerol-diphosphate]. Minor components contain unsubstituted 1,3-glycerol-diphosphate (chemical shifts within parentheses) as well as 1,3-glycerol-diphosphate with β-glucosyl and (2-acetamido-2-deoxy-β-glucopyranosyl) substituents.

4 Ribitol and glycerol are phosphorylated in position 1,5 and 1,3 respectively

Table 5 – ^{13}C teichoic acids chemical shifts (in ppm) from different bacterial species at 283K, pH 7.5 in 50mM HEPES.

α-GlcNAc				
	C1	95.1		
	C2	49.5		
	C3	78.2		
	C4	69.0		
	C5	72.7		
	C6	61.4		
	NHAc CO	175.2		
	NHAc CH_3	23		
β-Glc				
	C1	105.0		104.8
	C2	73.8		73.6
	C3			76.3
	C4			69.6
	C5			75.1
	C6			64.9
α-Glc				
	C1	98.5		
	C2	72.5		
	C3	74.0		
	C4	70.7		
	C5	72.9		
	C6	61.4		
β-GalNAc				
	C1			101.9
	C2			51.4
	C3			75.5
	C4			64.0
	C5			74.2
	C6			65.2
	NHAc CO			175.4
	NHAc CH_3			23.0

Table 5 – ^{13}C teichoic acids chemical shifts (in ppm) from different bacterial species at 283K, pH 7.5 in 50mM HEPES.

α-GalNAc				
	C1			94.1
	C2			50.2
	C3			67.6
	C4			77.3
	C5			71.3
	C6			64.3
	NHAc CO			175.4
	NHAc CH_3			23.0
AAT-Gal				
	C1			98.9
	C2			49.0
	C3			75.8
	C4			55.5
	C5			64.1
	C6			16.3

D

T1 TEMPERATURE DEPENDENCE

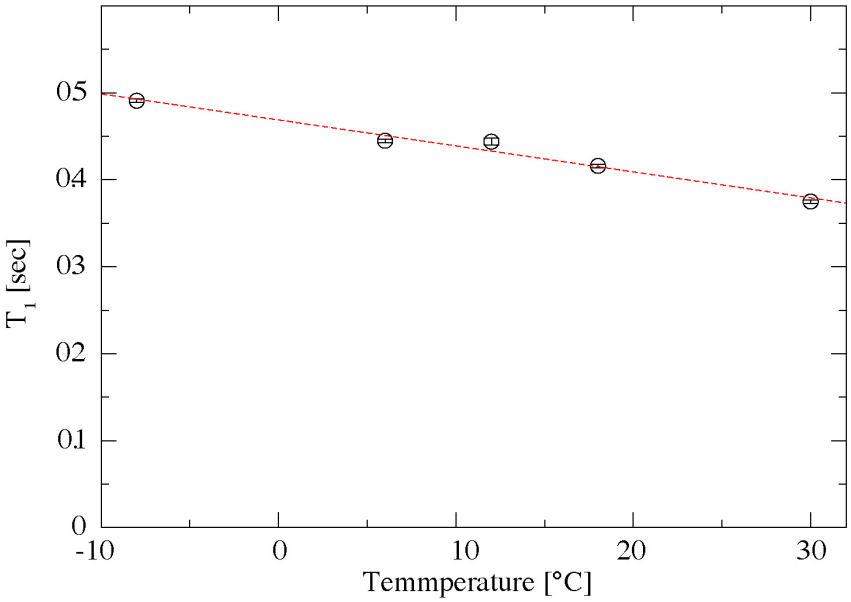


Figure 72 – T_1 temperature dependence of the ^{31}P resonance in *B. subtilis* 168. T_1 values have been measured on a 600MHz Varian spectrometer equipped with a 3.2 mm triple resonance BioMAS probe. An average T_1 time constant has been fitted from the integral over the entire spectrum. The spectra have been taken at a MAS frequency of 12.5 kHz and with an interscan delay of 5 s. Direct excitation has been used to create initial magnetization. The red line shows a linear fit of the values to represent the tendency of the T_1 to decrease with increasing temperature and flexibility.

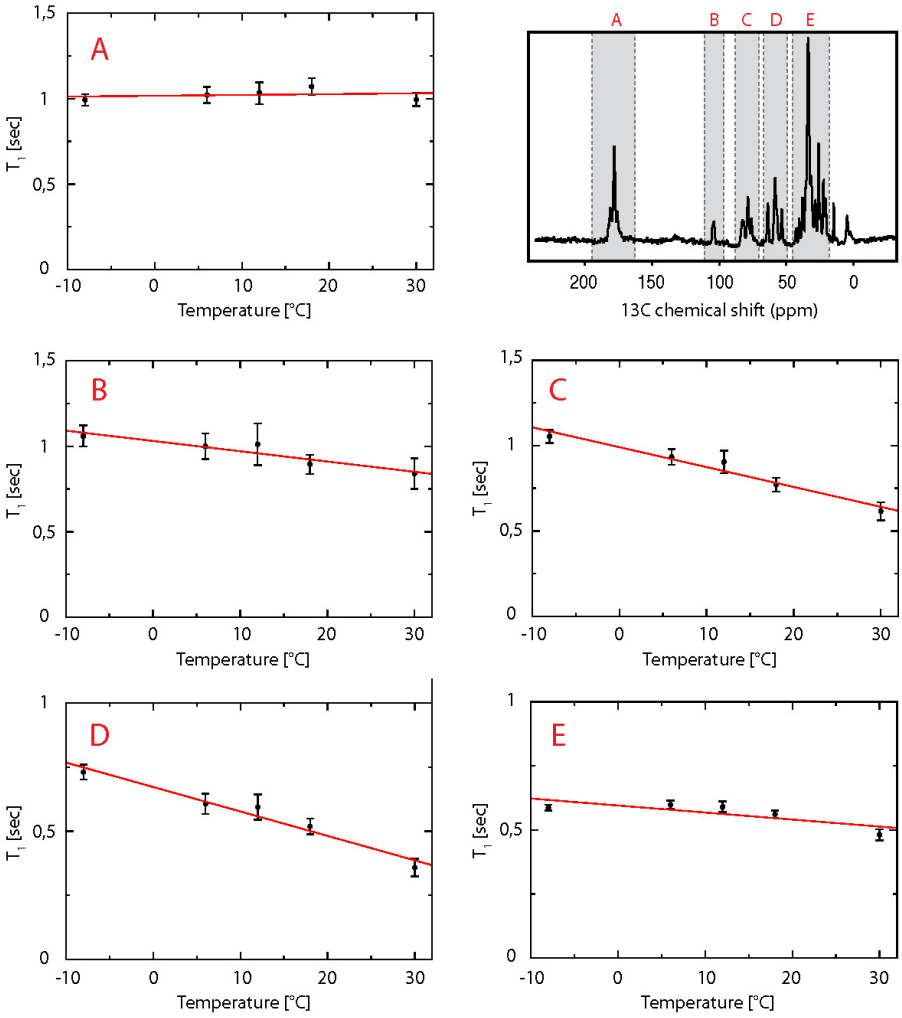


Figure 73 – T_1 temperature dependence of the ^{13}C resonances in *B. subtilis* 168 shown for different regions (indicated as A,B,C,D and E) . T_1 values have been measured on a 600MHz Varian spectrometer equipped with a 3.2 mm triple resonance BioMAS probe. The spectra have been taken at a MAS frequency of 12.5 kHz and with an interscan delay of 5 s. Direct excitation has been used to create initial magnetization. The red line shows a linear fit of the values to represent the tendency of the T_1 to decrease with increasing temperature and flexibility.

JBC Papers in Press. Published on July 27, 2009 as Manuscript M109.019471
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.M109.019471

THE CENTRAL DOMAIN OF DIVIB CAPS THE C-TERMINAL REGIONS OF THE FTSL/DIVIC COILED-COIL ROD

Soizic Masson[1], Thomas Kern[2], Audrey Le Gouëllec[1], Cécile Giustini[2], Jean-Pierre Simorre[2], Philip Callow[4], Thierry Vernet[1*], Frank Gabel[3] and André Zapun[1]

From [1]Laboratoire d'Ingénierie des Macromolécules, [2]Laboratoire de Résonance Magnétique Nucléaire, [3]Laboratoire de Biophysique Moléculaire, Institut de Biologie Structurale (UMR 5075, Université Joseph Fourier, CNRS, CEA), Grenoble, France,
and [4]EPSAM and ISTM Research Institutes, Keele University, Staffordshire, UK.

Running head: Structure of the DivIB/FtsL/DivIC complex

Address correspondence to: Thierry Vernet, Institut de Biologie Structurale, 41 rue Jules Horowitz, 38027 Grenoble, France. Tel.: +33 4 38 78 96 81 ; Fax: +33 4 38 78 54 94 ; e-mail : thierry.vernet@ibs.fr.

DivIB(FtsQ), FtsL and DivIC(FtsB) are enigmatic membrane proteins that are central to the process of bacterial cell division. DivIB(FtsQ) is dispensable in specific conditions in some species, and appears to be absent in other bacterial species. The presence of FtsL and DivIC(FtsB) appears to be conserved despite very low sequence conservation. The three proteins form a complex at the division site, FtsL and DivIC(FtsB) being associated through their extracellular coiled-coil region. We report here structural investigations by NMR, small-angle neutron and X-ray scattering, and interaction studies by surface plasmon resonance, of the complex of DivIB, FtsL and DivIC from *Streptococcus pneumoniae*, using soluble truncated forms of the proteins. We found that one side of the "bean"-shaped central β-domain of DivIB interacts with the C-terminal regions of the dimer of FtsL and DivIC. This finding is corroborated by sequence comparisons across bacterial genomes. Indeed, DivIB is absent from species with shorter FtsL and DivIC proteins that have an extracellular domain consisting only of the coiled-coil segment without C-terminal conserved regions (Campylobacterales). We propose that the main role of the interaction of DivIB with FtsL and DivIC is to help the formation, or to stabilize, the coiled-coil of the latter proteins. The coiled-coil of FtsL and DivIC, itself or with the transmembrane regions, could be free to interact with other partners.

Cell division is one of the defining features of life. Understanding the division of bacteria is also required in order to find novel antibiotic strategies. Numerous studies, carried out mostly with the model organisms *Escherichia coli* and *Bacillus subtilis* have uncovered several components of the divisome, which can be defined as the ensemble of proteins localized at the division site and participating to the process. Comparison of genomes and deletion studies indicate that the core of the divisome comprises eight conserved, mostly essential proteins: FtsZ, FtsA, FtsK, FtsQ(DivIB), FtsL, FtsB(DivIC), FtsW and FtsI. Fts nomenclature applies to Gram-negative organisms, whereas Div nomenclature applies to Gram-positive bacteria. These proteins are listed here in the conditional order of their recruitment to the division site of *E. coli* (1-4).

Processes in which they participate, have been attributed to several division proteins. FtsZ forms polymers with an annular distribution on the cytoplasmic side of the membrane and governs the recruitment of the other proteins. FtsA may mediate the interaction of FtsZ with the membrane. FtsK participates to the resolution of chromosome dimers, and possibly to the membrane fission. FtsI, and likely FtsW, participate to septal cell wall formation (1-4). In contrast, the roles of FtsQ(DivIB), FtsL, and FtsB(DivIC) have not been firmly linked to any particular process.

FtsQ(DivIB), FtsL and FtsB(DivIC) are positioned in the middle of the conditional order of recruitment in *E. coli* and *B. subtilis*. When the temporality of the recruitment was examined, FtsQ(DivIB) was found to belong to the late recruits, together with the proteins involved in cell wall assembly (5). In *E. coli*, the presence of FtsL and FtsB at the division site is mutually dependant, and their localization depends on that

of FtsQ (6,7). In *B. subtilis*, the presence of FtsL and DivIC at mid-cell depends on that of DivIB, at the temperature at which DivIB is essential, and reciprocally (8,9). A complex comprising FtsQ, FtsL and FtsB was isolated from *E. coli* by co-immunoprecipitation (10), and reconstituted *in vitro* with recombinant soluble forms of pneumococcal DivIB, FtsL and DivIC (11). The interaction of the three proteins was also confirmed by yeast and bacterial triple hybrid (12,13).

The genes *ftsL* and *ftsB*(*divIC*) are essential in *E. coli* and *B. subtilis* (6,14-16) and presumably in *Streptococcus pneumoniae* (17). The essentiality of *ftsQ*(*divIB*) in laboratory conditions varies between species. The gene *ftsQ* is essential in *E. coli* (18), but *divIB* is essential only at high temperature in *B. subtilis* (9,19), or in a chemically defined medium in *S. pneumoniae* (17). Under these conditions, the essentiality of DivIB appears to be a consequence of the protection from proteolysis that it provides to FtsL (8,17).

FtsQ(DivIB), FtsL and FtsB(DivIC) are bitopic membrane proteins with an N-terminal cytoplasmic region, a single transmembrane segment, and an extra-cytoplasmic region. The extracellular part is necessary and sufficient for the localization and function of FtsQ(DivIB), provided that it is anchored to the membrane (*e. g.* (20,21)), although the transmembrane segment also contributes to the localization (22,23). The extracellular part is organized in three regions termed α, β and γ. The crystal structure of a region consisting of the α- and β-domains was solved for FtsQ from *E. coli* and *Yersinia enterolitica* (24). The α-domain, comprising about 70 amino acids proximal to the cytoplasmic membrane, corresponds to the POTRA (for polypeptide transport-associated) domain first identified by sequence analysis and proposed to function as a molecular chaperone (25). The α- and β-domains form the conserved region of the FtsQ(DivIB) protein. The γ-region constitutes a C-terminal tail. It is highly variable in length and sequence and predicted to be unfolded. The γ-region was not observed in the structures from *E. coli* and *Y. enterolitica*, thus confirming its flexible nature (24).

The α-domain in the recombinant soluble form of the extracellular part of DivIB from *Geobacillus stearothermophilus* was digested by trypsin and therefore considered to be largely unfolded (26). The γ-region was also removed by trypsin digestion, together with a C-terminal fragment of the β-domain. The structure of the resulting shorter β-domain from *G. stearothermophilus* was solved by NMR (26) and lacks the two C-terminal β-strands.

Localization epitopes have been identified in the transmembrane segment, the α-domain and a region encompassing the C-terminal part of the β-domain and γ-tail of DivIB from *B. subtilis* (23). Likewise in *E. coli*, a region in the α-domain is required for the localization of FtsQ, whereas the C-terminal region of the β-domain and the last α-helix are required for the recruitment of FtsL and FtsB (24). In *S. pneumoniae*, the essentiality of DivIB in defined medium was found to reside in the C-terminal region of the β-domain (17).

No experimental structure is known for FtsL or FtsB(DivIC). Both are small proteins comprising between 90 and 140 amino acids. The number of residues is sometimes larger, as in *Mycobacterium tuberculosis* (384 for FtsL and 228 for FtsB), due to N- and/or C-terminal extensions consisting of mostly charged and polar amino acids or proline-rich sequences. The major part of FtsL or FtsB(DivIC) is extracellular and contains a region, proximal to the transmembrane segment, predicted to form a coiled-coil of about five heptads. Coiled-coil helices associate longitudinally to mediate protein association. It is possible that the coiled-coil helices are continuations of the transmembrane helices, although a proline (known to break helices) is present in some species between the two segments. Following the coiled-coil region is a 25-35 residue long C-terminal region in both FtsL and DivIC(FtsB). This region was recently shown in FtsB to be required for the interaction with FtsQ in *E. coli* (27).

We report here the results of structural studies in solution of a ternary complex consisting of the β- and γ-segments of DivIB, and a constrained dimer of the extracellular parts of FtsL and DivIC from *S. pneumoniae*. Despite the coiled-coil predictions, the recombinant extracellular domains of FtsL and DivIC did not interact *in vitro* (11,28). Forced dimerization was obtained by fusion with the artificial coiled-coil

peptides k5 and e5 (35 residue-long), which are known to form a heterodimer due to their complementarity of charge, with a nanomolar dissociation constant (29). The k5- and e5-coils were fused to the extracellular domain of FtsL and DivIC, to give rise to KL and EC fusion proteins, respectively. The constrained dimer (KL/EC) was shown to interact with the extracellular part of DivIB ($DivIB_{ext}$), yielding a soluble complex amenable to structural studies (11).

The overall shape of the complex and its constituents was probed using X-ray and neutron small-angle scattering (SAXS and SANS). NMR was used to investigate the interface between the proteins by chemical shift mapping. The interaction was further investigated using surface plasmon resonance with truncated forms of the proteins. The complex of DivIB, FtsL and DivIC is formed by the interaction of one face of the β-domain of DivIB with the C-terminal regions of FtsL and DivIC, at the tip of an elongated rod formed by the coiled-coil segments. The α-domain of DivIB and the coiled-coil regions of FtsL and DivIC remain free to interact with other proteins of the division apparatus.

EXPERIMENTAL PROCEDURES

Plasmid constructions- Proteins were produced in *E. coli* from expression plasmids inducible with isopropyl-β-D-1-thio-galactopyranoside. The sequence encoding $DivIB_{ext}$ preceded by the sequence for a Tobacco Etched Virus protease cleavage site (TEV) was introduced into pGEX-4T1 (GE Healthcare) to produce the fusion protein GST-TEV-$DivIB_{ext}$ (11). The sequence encoding the βγ-fragment (residues 224 to 396) was similarly introduced into pGEX-4T1 to produce GST-TEV-βγ. The sequence encoding the β-domain (residues 224 to 361) following a TEV site was introduced into pLX06 (Protein'eXpert) to produce the fusion protein GST-TEV-β.

To produce the KL protein (residues 44 to 105 of FtsL fused following the k5 peptide) with a N-terminal poly-His-tag, and a tryptophan to generate absorption at 280 nm, the codon TGG was introduced by site-directed mutagenesis into the pETduet-based (Novagen) plasmid described previously (11). To produce KL* (as KL but ending at residue 82 of FtsL), the stop codon TAG was introduced by site-directed mutagenesis. To produce EC (residues 55 to 122 of DivIC fused following the e5 peptide) with a N-terminal Strep-tag (IBA), the sequence encoding Strep-EC was amplified from the plasmid described previously encoding a poly-His-tagged EC (11) and introduced in pET30 (Novagen). The sequence encoding a poly-His-tag followed by a TEV site and EC was similarly introduced into pET30 (Novagen), whereas the sequence encoding a poly-His-tag followed by a TEV site and EC* (as EC but ending at residue 93) was introduced into pLIM09 (RoBioMol). A plasmid was also constructed for the co-expression of KL and EC by sub-cloning the sequence encoding the Strep-tagged EC into the pETduet-based plasmid encoding the poly-His-tagged Trp-containing KL.

Protein production and purification- Unlabeled proteins were expressed in *E. coli* BL21(DE3) RIL in flasks or in a fermentor (Minifors, Infors) in Luria Broth or Terrific Broth and the appropriate antibiotic at 37 °C following the addition of 0.5 mM isopropyl-β-D-1-thio-galactopyranoside. For the production of ^{15}N-labeled $DivIB_{ext}$, cells were grown in a fermentor at 28 °C in minimal medium (KH_2PO_4 3 g/L; Na_2HPO_4 6 g/L; NaCl 0.5 g/L; $MgSO_4$ 1 mM; $FeCl_3$ 50 μM; $CaCl_2$ 20 μM; $MnCl_2$ 10 μM; $ZnSO_4$ 10 μM; $CoCl_2$ 2 μM; $CuCl_2$ 2 μM; $NiCl_2$ μM; Na_2MoO_4 2 μM; H_3BO_3 2 μM; biotin 1 μM; niacinamide 5 μM; pyridoxine 5 μM; thiamine 5 μM; folic acid 1 μM; riboflavin 1 μM; glucose 4 g/L) with $^{15}NH_4Cl$ 1 g/L as nitrogen source. Induction was with 1 mM isopropyl-β-D-1-thio-galactopyranoside for 3 h at 30 °C. For the production of $^{15}N/^{2}H$-labeled β-domain, cells were grown in minimal medium as above in D_2O with $^{15}NH_4Cl$ 1 g/L as nitrogen source. For the production of $^{15}N/^{2}H/^{13}C$-labeled β-domain, cells were grown as above in D_2O with $^{15}NH_4Cl$ 1 g/L and ^{13}C-glucose 4 g/L as nitrogen and carbon source.

$DivIB_{ext}$, β and βγ were purified in a first step by glutathione affinity chromatography. Cell lysates in 20 mM Tris pH 8, 500 mM NaCl, containing the GST-TEV-$DivIB_{ext}$, GST-TEV-β or GST-TEV-βγ fusion proteins were loaded onto glutathione Sepharose resin (GE Healthcare). The fusion proteins were cleaved on the resin by

incubation at 37 °C with the poly-His-tagged TEV protease and recovered by washing with 20 mM Tris pH 8, 150 mM NaCl. The TEV protease was then removed by binding on Ni-NTA resin (Qiagen). This procedure was sufficient for the ^{15}N-DivIB$_{ext}$ protein used for recording HSQC [^{15}N;^{1}H] spectra. ^{15}N/^{2}H- and ^{15}N/^{2}H/^{13}C-labeled β-domains were further purified by size exclusion chromatography on a Superdex S75 column (16 x 600 mm) equilibrated with 20 mM Tris pH 8, 150 mM, or 20 mM sodium phosphate pH 6, 150 mM NaCl.

Further purification steps were also applied before scattering experiments. DivIB$_{ext}$, β and βγ proteins from the affinity purification were dialyzed against 20 mM Tris pH 8 to lower the NaCl concentration to 15 mM, prior to loading onto a Resource Q resin (GE Healthcare) and elution with a NaCl gradient from 15 to 500 mM in ten column volumes. Fractions of interest were pooled and concentrated with Amicon Ultra devices (Millipore) prior to size exclusion chromatography on a Superdex 75 column (10 x 300 mm) (GE Healthcare). Homogenous fractions (as judged by SDS- and non denaturing polyacrylamide gel electrophoresis) were pooled, dialyzed against 20 mM Tris pH 8, 100 mM NaCl, and finally concentrated with Amicon Ultra devices.

The protein EC with a Strep-tag was purified from the cell lysate in 20 mM Tris pH 8 150 mM NaCl by affinity chromatography on a Streptactin resin (IBA). The protein was eluted with 2.5 mM desthiobiotine in the same buffer.

The proteins EC, KL, EC* and KL* with poly-His-tags were purified from the cell lysates in 20 mM Tris pH 8, 500 mM NaCl, 20 mM imidazole, by affinity chromatography on Ni-NTA resin (Qiagen). Proteins were eluted with an imidazole gradient of 20 to 500 mM in 10 column volumes.

The (KL/EC), (KL*/EC) and (KL/EC*) complexes, with both partners harboring poly-His-tags, were prepared from proteins purified as described above. The two proteins were mixed in equivalent amount. Urea was added in order to reach a concentration of 8 M. The urea was then removed by dialysis three times against 40 volumes of 20 mM Tris pH 8, 150 mM NaCl.

The (KL/EC) complex, where EC has a Strep-tag and KL a poly-His-tag, was prepared from a lysate of cells co-expressing both proteins. The complex was isolated by two successive steps of affinity chromatography in 20 mM Tris pH 8, 150 mM NaCl, first on Ni-NTA resin eluted with an imidazole gradient from 20 to 500 mM, then on Streptactin eluted with 2.5 mM desthiobiotine. Following concentration with an Amicon Ultra devices, the complex was further purified by size exclusion chromatography on a Superdex S200 column (16 x 600 mm) equilibrated with 20 mM Tris pH 8, 150 mM. For the TROSY NMR experiment, the buffer was exchanged by desalting on a PD10 column (GE healthcare) against 100 mM ammonium acetate pH 7 prior to lyophilization.

Determination of protein concentrations- Protein concentrations were determined by the absorbance at 280 nm using the following theoretical extinction coefficient: 5500 $M^{-1}cm^{-1}$ for poly-His-tagged KL and KL*, 17420 $M^{-1}cm^{-1}$ for Strep-EC, 13410 $M^{-1}cm^{-1}$ for poly-His-tagged EC, 4470 $M^{-1}cm^{-1}$ for poly-His-tagged EC*, 26360 $M^{-1}cm^{-1}$ for DivIB$_{ext}$, 19858 $M^{-1}cm^{-1}$ for βγ and 11920 $M^{-1}cm^{-1}$ for β. The absorbance spectra were identical without and with 8 M urea indicating that the calculated extinction coefficient could be used without correction. For the KL/EC complex (purified after coexpression of poly-His-tagged KL and Strep-EC) an amino acid analysis gave an experimental extinction coefficient of 24740 $M^{-1}cm^{-1}$, which is 8% larger than the calculated extinction coefficient (22920 $M^{-1}cm^{-1}$); we used the theoretical coefficient.

NMR experiments- Exploratory heteronuclear NMR data were collected at 27°C in a 20 mM Tris buffer (90% (v/v) H_2O, 10% (v/v) D_2O) at pH 8 with 150 mM NaCl, on samples of DivIB$_{ext}$ at 140 μM, DivIB$_{ext}$(KL/EC) at 110 μM and β at 270 μM. Spectra were recorded on a 800 MHz Varian spectrometer equipped with a cold probe. For backbone assignment, a sample of uniformly ^{13}C/^{15}N-labeled β-domain at 0.8 μM was prepared and data were collected at 27°C in a 20 mM sodium phosphate buffer (90% (v/v) H_2O, 10% (v/v) D_2O) at pH 6 with 150 mM NaCl. A set of high field triple resonance experiments (HNCO, HNCαCO, HNCαCβ, CβCαCONH) where acquired and assignment was done using the CO, C_α and C_β resonances. First assignment

was obtained by an automatic approach using MARS (30) and was then refined manually.

The binding of (KL/EC) to the β-domain was studied by TROSY-type ^{1}H-^{15}N correlation experiments. NMR data were collected at 25°C with 32 transients and 1024 (^{1}H) x 64 (^{15}N) complex points. A reference spectrum was taken on a 0.4 mM uniformly $^{13}C/^{15}N$-labeled deuterated β-domain sample and lyophilized (KL/EC) was added to a concentration of 0.5 mM.

SAXS and SANS experiments and data processing- The scattered intensities $I(Q)$ were analyzed as a function of the wave vector transfer Q, with equation 1: $Q = (4\pi/\lambda)\sin\theta$, where 2θ is the scattering angle and λ the X-ray or neutron wavelength.

The SAXS measurements were performed on the ID02 beamline (31) at the European Synchrotron Radiation Facility (Grenoble, France) at a sample-to-detector distance of 2 m to cover Q-range from 0.008 to 0.28 Å^{-1}. The X-ray wavelength was 1 Å and data were collected by a high sensitivity CCD (FReLoN) detector placed in an evacuated flight tube. Solutions were loaded in a flow-through quartz capillary cell (diameter 2 mm; wall thickness 10 µm) at 15°C. The radiation damage was checked with 10 successive exposure times of 0.1 s each on a single part of the sample. Longer exposure time led to damage. Therefore, we acquired and averaged 10 curves of 0.1 s exposure time on different parts of the sample that was pushed through the capillary cell. The 2D diffraction patterns were normalized to an absolute scale and azimuthally averaged to obtain the intensity profiles $I(Q)$, and the solvent background was subtracted using Fit2D (32).

SANS was carried out on the large dynamic range small angle diffractometer D22 (33) at the Institut Laue-Langevin (Grenoble, France). Samples of 200 µl in H_2O buffers were measured in Hellma® QS 1 mm quartz cells at 4°C. (Data sets from samples in D_2O could not be used due to aggregation problems.) The sample-to-detector distances were 2 and 8 m with an incident neutron wavelength of 6 Å to cover a Q-range from 0.025 to 0.30 Å^{-1}. The sample and buffer raw 2D intensities (after masking inappropriate data points) were normalized to the incoming neutron flux and corrected for detector efficiency (H_2O reference), electronic background (B_4C standard) and sample holder scattering (empty quartz cuvette), by using RNILS (34). Corrected 2D sample and buffer intensities were summed up azimuthally into 1D $I(Q)$ intensities using SPOLLY (34). Buffer intensities were then subtracted using PRIMUS (35).

Determination of molecular weights- Molecular masses M_r were determined by SAXS relatively to hen egg white lysozyme using equation 2: $M_r = I(0)/f_2C$, where $I(0)$ is the forward scattering and C is the protein concentration in g/L. The factor $f_2 = (83\pm1)10^{-6}$ molLg^{-2} was determined using hen egg white lysozyme at a concentration of 10.0 g/L (36).

Molecular weights M_r were determined on an absolute scale by SANS using equation 3 (37):

$$M_r = \frac{1-T}{f4\pi T_S}\frac{I(0)}{I_{inc}(0)}\frac{10^3}{CN_A t}\left[\left(\sum b_i - \rho_S V\right)/M_r\right]^{-2},$$

where $I(0)$ is the coherent macromolecular forward scattering (extracted by the Guinier approximation, equation 4), $I_{inc}(0)$ is the incoherent forward scattering from H_2O, T_s and T are the transmissions of the sample and H_2O, respectively. C is the protein concentration in g/L, t is the sample thickness in cm, f is a correction factor for the anisotropicity of the solvent scattering as a function of neutron wavelength, b_i are the scattering lengths of the protein atoms in cm, ρ_s is the solvent scattering density in cm^{-2}, V is the protein excluded volume and N_A is the Avogadro constant.

Determination of the radii of gyration- We extracted the radii of gyration R_g of all samples by a Guinier analysis using equation 4 (38): $\ln[I(Q)] = \ln[I(0)] - \frac{1}{3}R_g^2Q^2$, where $I(0)$ is the intensity scattered in the forward direction. The $I(0)$ and R_g values were extracted by the Guinier approximation using the program PRIMUS (35). The validity of the approximation was checked *a posteriori* in each case and was always fulfilled ($R_gQ_{(max)}<1$).

The cross-sectional radius of gyration R_C was determined according to equation 5:

$$I_C(Q) = I_C(0)\exp\left(-\frac{1}{2}Q^2R_C^2\right) \text{ with } I_C(Q) = I(Q)Q/\pi L,$$

with a graphical fit using the program PRIMUS (35). The good agreement between the

radii of gyration and the maximum extension of the particle ($L = D_{max}$) was checked using the relationship (equation 6) $\frac{L^2}{12} = R_G^2 - R_C^2$.

Distance distribution functions and ab initio *modeling*- The distance distribution functions $P(r)$ were determined by the program GNOM (39) from SAXS and SANS curves. The boundary conditions $P(0) = P(D_{max}) = 0$ were imposed in each case, D_{max} being the maximum extension of the particles. Low-resolution envelopes were calculated from the $P(r)$ functions with DAMMIN (40).

Rigid body modeling- The experimental SANS data from the complex βγ(KL/EC) were modeled using the βγ *ab initio* model and the coiled coil structure of cortexilin I (PDB entry 1D7M) as rigid bodies using the program MASSHA with default parameters (41). In addition, several arrangements of both rigid bodies were generated manually using MOLMOL (42) and scored against the SANS curve using CRYSON (43).

Calculation of scattering intensities from known structures- Theoretical scattering curves $I(Q)$ of the *E. coli* and *G. stearothermophilus* β-domains were calculated from the PDB entries 2VH1 and 1YR1 using the program CRYSOL (44) with default parameters.

Surface plasmon resonance measurements- Measurements were performed with a Biacore™ 3000 instrument. All samples were dialyzed against 10 mM Hepes pH 7.5, 150 mM NaCl, 50 mM EDTA and 0.005% P20. The dialysis buffer was used as running solution throughout. KL/EC, KL*/EC and KL/EC* mixtures (with all proteins carrying an N-terminal poly-His-tag) were attached to a NTA sensor chip activated with 10 µL of $NiCl_2$ 500 µM. Experiments were at room temperature with a 20 µL/min flow rate. The (KL/EC), (KL*/EC) and (KL/EC*) dimers (ligands) were injected at the concentration of 100 nM. The $DivIB_{ext}$, β and βγ proteins (analytes) were injected at concentration ranging from 5 µM to 1.5 mM. For each experiment, interaction curves recorded for the running buffer on the ligand, and for the analyte on the naked sensor chip, were subtracted. Sensorgrammes were analyzed with the program BIAevaluation 3.0.

RESULTS

Sequence comparisons- The coiled-coil regions of FtsL and DivIC(FtsB) are usually followed by 25 to 30 amino acids that are C-terminal, except in species with long additional unfolded extensions. There is little sequence conservation between species, and it is sometimes difficult to identify the genes encoding FtsL or DivIC(FtsB) using sequence comparison methods. Instead, it is best to look for genes encoding proteins with a single transmembrane segment followed by a predicted coiled-coil region, in the proper chromosomal environment. The *ftsL* gene is usually following *mraW*, while *divIC*(*ftsB*) is usually close to that of enolase.

A set of FtsL and DivIC(FtsB) sequences were aligned using the program CLUSTALW. The transmembrane region was predicted using the whole alignments with the program TMAP. The coiled-coil regions were predicted using PAIRCOIL2. The alignments are presented in supplementary material Figs S1 and S2.

Despite the absence of overall sequence conservation of FtsL, three hydrophobic amino acids (a Val/Leu/Ile, an Ala, and a Leu/Met) are conserved in the region following the predicted coiled-coil (Fig. S1). In FtsB(DivIC), a basic (Arg) and an acidic (Glu) residue are conserved in the region following the coiled-coil (Fig. S2).

It is noteworthy that the extracellular region of both FtsL and FtsB are shorter in Campylobtacterales, with very little sequence following the predicted coiled-coil region, and are lacking the conserved residues mentioned above. Remarkably, Campylobacterales appear to be also devoid of FtsQ, suggesting that the C-terminal regions of FtsL and FtsB(DivIC) containing the conserved residues might mediate the interation with (FtsQ)DivIB.

Limited proteolysis- In order to probe the domain organization of DivIB from *S. pneumoniae*, we submitted $DivIB_{ext}$ (corresponding to residues 150 to 396) to trypsin digestion. A time course monitored by Coomassie-stained SDS-polyacrylamide gel electrophoresis is shown in Fig. 1. $DivIB_{ext}$ (28177 Da) migrated as a 28 kDa protein. After 50 min of digestion with trypsin, a major stable fragment migrated as a 15 kDa protein. The mass

was determined by electrospray mass spectrometry to be 15842±1 Da. N-terminal sequencing yielded the sequence VKEYDIVA. Combining the mass and sequence data indicated that the fragment spans residues 220 to 361 of pneumococcal DivIB (15843 Da). This fragment overlaps the β-domain defined by the structure of FtsQ from *E. coli*, and is somewhat longer on the C-terminal side than the β-domain defined by partial proteolysis of DivIB from *G. stearothermophilus*. The trypsin-sensitive region 150-219 and 362-396 would correspond to regions α and γ, respectively.

For further studies, two truncated forms of pneumococcal DivIB that we termed β and βγ, corresponding to residues 224 to 361, and 224 to the C-terminus 396, respectively, were purified. We decided to produce proteins with a truncated N-terminus with respect to our proteolysis results in order to exclude the fully conserved E222, that is absent from the β-domain from *G. stearothermophilus*, and appears to be part of the linker between the α- and β-domains from *E. coli* and *Y. enterocolitica*. Both β and βγ contain an additional N-terminal glycine resulting from the cleavage of the fusion with the glutathione-S-transferase (GST) using the TEV protease.

NMR investigations of the extracellular part of DivIB and its interaction with a soluble artificially constrained dimer of FtsL and DivIC- To gather information on the interaction between DivIB and the (FtsL/DivIC) dimer, we undertook NMR studies of the soluble model systems consisting of the complete extracellular $DivIB_{ext}$ or the β-domain, with or without the presence of the soluble dimer (KL/EC). The [^{15}N-^{1}H] HSQC spectrum of ^{15}N-labeled $DivIB_{ext}$ showed about a hundred well-spread isolated peaks, as well as an area of unresolved peaks in the vicinity of 120 ppm; 8 ppm (corresponding to the ^{15}N and ^{1}H dimensions, respectively) (not shown). [^{15}N-^{1}H] HSQC spectra show resonance peaks corresponding to every proton attached to ^{15}N atoms, *i. e.* the amide protons of the backbone chain and the few amine protons from side chains. Amide protons from unstructured polypeptides present a low dispersion of the resonances centered for the proton around 8 ppm. Amide protons in folded structures are each in particular environments that modify their resonance frequencies (chemical shift) resulting in the isolation and spreading of their peaks. The spectrum therefore indicated that the 247-residue-long $DivIB_{ext}$ contains a folded domain of about 100 amino acids, and about 150 unstructured residues.

The [^{15}N-^{1}H] HSQC spectrum of the β-domain showed about 100 well dispersed peaks super-imposable to those of $DivIB_{ext}$ (not shown), without an accumulation of resonances in the 120 ppm; 8 ppm region. The comparison of the spectra of $DivIB_{ext}$ and β showed that the trypsin-sensitive α- and γ-segments were largely unstructured.

When an excess of unlabeled (KL/EC) was added to ^{15}N-$DivIB_{ext}$, at a concentration where most of the $DivIB_{ext}$ should be interacting with the dimer, the [^{15}N-^{1}H] HSQC spectrum showed the disappearance of nearly all the dispersed peak and the persistence of the peaks due to unfolded parts of the protein (not shown). We interpreted these observations as indicating that the (KL/EC) dimer interacts with the β-domain of DivIB, without important contact with the α- or γ-regions. When the same experiment was performed with only the ^{15}N-β-domain, all the peaks disappeared. The disappearance of the peaks from the β-domain upon association with (KL/EC) could be explained by a greater transverse relaxation resulting from a change in the hydrodynamic behavior of the protein. The β(KL/EC) complex likely assumes a non-compact shape that drastically reduces its rotational diffusion. If the complex were compact, the two-fold difference in size between the free β-domain and β(KL/EC) would not reduce the rotational diffusion to the point of extinguishing the signal. The persistence of the signal of the α- and γ-regions observed on the complete ^{15}N-$DivIB_{ext}$ in presence of (KL/EC) indicates that these do not participate in the interaction, and that they retain their flexibility.

The disappearance of the signal upon formation of the ternary complex precluded further use of the ^{15}N-labeled samples. To minimize the translational relaxation, and thus compensate the loss of signal due to the loss of rotational diffusion upon formation of the ternary complex, we prepared a ^{15}N/^{2}H doubly labeled sample of β. The overall deuteration reduces the translational relaxation, whereas atoms that can be exchanged with ^{1}H protons from the solvent, such as amide

protons, remain visible in [^{15}N-^{1}H] HSQC experiments.

Indeed, the [^{15}N-^{1}H] HSQC spectrum of the ^{15}N/^{2}H-β-domain was identical to that of the ^{15}N-labeled sample. In contrast to what had been observed with the ^{15}N-labeled β-domain, it was still possible to record a good quality spectrum following addition of an excess of (KL/EC) to the ^{15}N/^{2}H-β-domain (Fig. 2A). The spectrum of the free β-domain exhibited 126 resolved peaks at 27°C. The spectrum of β in the presence of (KL/EC) revealed a modification of chemical shift for 27 peaks. The modification of the chemical shift of some protons of the β-domain indicates that their local environment is changed in the complex with (KL/EC). To identify the position of these protons, it was necessary to assign the peaks to specific protons of the protein. For this purpose, we produced a doubly labeled ^{13}C/^{15}N-sample of the β-domain. Assignment of the ^{1}H$_{N}$ resonance peaks was achieved by analyzing HSQC, HNCO, HNCACO, HNCACB, HNCA and CBCACONH spectra. Of the 126 peaks of the HSQC [^{15}N; ^{1}H] spectrum, 97 could be assigned, including 27 that experienced modification of their chemical shift upon binding of (KL/EC). The attribution is nearly complete for the first three quarters of the β-domain (residues 224 to 323) but could not be achieved for residues 324 to 352 (Fig. 2B), probably due to the dynamics of conformational exchanges of the region. Secondary structure propensities obtained from the ^{13}C$_{\alpha}$ and ^{13}C$_{\beta}$ chemical shifts of the pneumococcal protein, compared to the secondary structures of the *E. coli* and *G. stearothermophilus* proteins, are given as supplementary material (Fig. S3).

Small-angle X-ray scattering investigation of the extracellular part of DivIB and of (KL/EC)- To obtain low-resolution structural information on the extracellular part of DivIB in solution, a series of small-angle X-ray scattering measurements were performed on various truncated forms of the protein.

Small-angle scattering curves provide information on the mass and shape of the scattering particle. The mass is proportional to the intensity scattered in the forward direction ($I(0)$) provided that the sample is dilute and that there is a single scattering species, and can be extracted from SAXS and SANS data using equations 2 and 3, respectively (see Experimental procedures). The shape of the scattering curve $I(Q)$ is related to the shape of the particle and yields the radius of gyration, R_g, at the smallest scattering angles (Guinier approximation, equation 4, Experimental procedures). The distribution of distances between scattering points of the particle (termed distance distribution function $P(r)$) can be extracted from the whole scattering curve. Combined, the radius of gyration, the profile of the distance distribution function and the molecular mass hint at the shape of the particle (*e. g.* elongated *vs* compact). Theoretical distance distribution functions and scattering curves can be unambiguously calculated for any particle of known atomic structure. The reverse is not true, but three-dimensional models of a particle can be constructed by iterative numerical simulation.

The complete extracellular part DivIB$_{ext}$ produced scattering curves with very high forward scattering $I(0)$ (not shown), indicating that samples were not monodisperse but contained oligomers or aggregates. Masses determined from the $I(0)$ values increased with the protein concentration. The presence of aggregates of DivIB$_{ext}$ was also observed by analytical ultracentrifugation (not shown). The formation of aggregates was not unexpected given that DivIB$_{ext}$ is partly unfolded, but precluded the use of SAXS or SANS for further investigation.

The shape of the SAXS curves obtained for the β and βγ proteins showed no variation with the concentration (between 117 and 685 μM for β, between 33 and 313 μM for βγ). The apparent masses determined from the $I(0)$ values were consistent with the theoretical masses. The measured mass for β was 16±2 kDa for an expected mass M_{theor} = 15 kDa, that of βγ was 25±3 kDa for an expected mass M_{theor} = 20 kDa. The shape of the scattering curves and the good agreement between measured and calculated masses showed that the samples were largely monodisperse, as observed by analytical ultracentrifugation (not shown).

For comparison, theoretical scattering curves were calculated for the β-domain of DivIB from *G. stearothermophilus* (residues 115-233) and FtsQ from *E. coli* (residues 126-260) using the program CRYSOL (44). The normalized experimental and theoretical curves are shown in

Fig. 3A. The experimental SAXS curve of the β-domain from *S. pneumoniae* is very similar to that calculated for the *E. coli* orthologue, indicating that both protein domains have the same overall shape. However, the radii of gyration R_g of both β-domains are somewhat different (20.8±1 Å and 17.4 Å, for *S. pneumoniae* and *E. coli*, respectively), despite having nearly the same mass (15381 Da and 15317 Da). The pneumococcal β-domain is therefore likely more elongated, as shown by the distance distribution function $P(r)$ in Fig. 3B, which displays more long distances.

In contrast, the calculated scattering curve from the β-domain of *G. stearothermophilus* deviates significantly from those from *S. pneumoniae* and *E. coli* (Fig. 3A). At small Q-values, it decreases less rapidly, whereas the intensity decreases strongly at larger Q-values ($Q > 0.15$ Å^{-1}), suggesting that the β-domain from *G. stearothermophilus* is more compact than its pneumococcal orthologue, with a radius of gyration $R_g = 15.3$ Å.

The experimental SAXS curve of the βγ-fragment decreases more rapidly at small Q-values than those of the β-domains from *S. pneumoniae* and *E. coli*, indicating a more elongated conformation (Fig. 3A). At large Q-values ($Q > 0.13$ Å^{-1}) curves of β- and βγ-proteins look similar. The $P(r)$ of the βγ-fragment is nearly identical to that of the β-domain, with one shoulder at greater distances. This comparison suggests that the βγ-protein has the same shape as the β-domain with a longitudinal extension.

Ab initio models of the β- and βγ-proteins were calculated from the distance distribution functions. Five independent models were produced and were consistent (Figs 3C and supplementary material S4). The resolution of the β models is coarser than that of βγ, as the data were noisier at large Q-values. One representative model of each pneumococcal truncated form (β and βγ) is shown in Fig. 3C, alongside structures of the β-domain from *G. stearothermophilus* and *E. coli*. The β-domain of pneumococcal DivIB appears to be elongated and slightly curved, or "bean"-shaped. Models of the βγ protein fragment comprise a "bean"-shaped domain and an additional smaller globular domain attached by a short stalk at one extremity of the "bean", which likely constitutes the γ-tail. The γ-region is unstructured as judged by its sensitivity to trypsin and the NMR data. The smaller globular domain of the βγ model could represent the conformational sampling by the unfolded γ-tail.

As expected from the direct comparison of the SAXS and theoretical scattering curves, the pneumococcal β-domain has an overall shape similar to that of the *E. coli* β-domain. The crystal structure of the *E. coli* β-domain can be placed easily within the model envelope of the "bean"-shaped domain of the βγ-fragment (Fig. 3D).

Relative molecular mass measurements by SAXS of the (KL/EC) artificially constrained dimer at various concentrations indicated that the dimer was self-associating to form a tetramer $(KL/EC)_2$ with a dissociation constant comprised between 5 and 50 μM (not shown). We did not pursue the structural characterization of the $(KL/EC)_2$ species as it is likely an artefact resulting from the higher order association of the k5/e5 coiled-coil, which is known to form $(k5/e5)_2$ tetramers at high concentrations (29). Note that hydrodynamic techniques such as size exclusion chromatography, dynamic light scattering or analytical ultracentrifugation failed to reveal the formation of the tetramer $(KL/EC)_2$, presumably because the (KL/EC) and $(KL/EC)_2$ particles have comparable hydrodynamical properties despite having a two-fold mass difference, due to the elongated, highly anisotropic nature of the coiled-coils.

Small angle neutron and X-ray scattering investigations of the ternary complex βγ(KL/EC)- At high concentrations of βγ/KL/EC (> 10 μM), necessary to generate high quality scattering data, we could have expected the presence of species such as $\beta\gamma(KL/EC)_2$ or $\beta\gamma_2(KL/EC)_2$, of unknown biological significance. The formation of $(KL/EC)_2$ being unknown to us at the time, we nevertheless acquired SANS and SAXS curves of KL/EC mixtures, βγ alone, and βγ/KL/EC.

Aggregation in 42% and 100% D_2O prevented using the differential deuteration of subunits in various H_2O/D_2O solutions to identify the individual components within the complex, as initially planned. A SANS experiment in water is presented here and data are summarized in Table 1. The mass of the scattering species in the KL/EC sample (55±8 kDa) was consistent with

that of the tetramer $(KL/EC)_2$ (M_{theor} = 50 kDa), as expected from the SAXS results for the concentration of 270 μM. The measured mass of the βγ-fragment (22±5 kDa) was also consistent with that of a monomer (M_{theor} = 20 kDa). The mean mass of the scattering species in the sample containing stoichiometric amounts of KL/EC and βγ-fragment (46±6 kDa) did not correspond to that of a complex $\beta\gamma_2(KL/EC)_2$ (M_{theor} = 90 kDa), neither to that of a non interacting mixture of βγ and $(KL/EC)_2$ (mean M_{theor} = 35 kDa). Instead, the measured mass was consistent with that of the βγ(KL/EC) complex (M_{theor} = 45 kDa), or possibly with a mixture of $\beta\gamma(KL/EC)_2$ and free βγ (mean M_{theor} = 45 kDa). Analytical ultra-centrifugation of the sample showed the presence of a single species (not shown), indicating that the acquired SANS curve was likely that of the ternary βγ(KL/EC) complex.

The masses measured by SANS suggested that the association with the βγ-protein could drive the dissociation of the tetramer $(KL/EC)_2$ into the (KL/EC) dimer. The reaction could be described by the two competing equilibria, βγ + (KL/EC) ⇆ βγ(KL/EC) with the constant K_{D1}, and $(KL/EC)_2$ ⇆ 2(KL/EC) with the constant K_{D2}. As we had values of the equilibrium constants K_{D1} = 0.22 μM for the dissociation of the βγ(KL/EC) complex determined by surface plasmon resonance (see below), and K_{D2} comprised between 5 and 50 μM for the dissociation of $(KL/EC)_2$ estimated from SAXS mass measurements, we ran numerical simulations of the composition of the mixture at equimolar concentrations of βγ and (KL/EC). The simulations showed that for concentrations greater than 100 μM, the sample would contain mostly the βγ(KL/EC) complex, with negligible amounts of free species and $(KL/EC)_2$ tetramer (supplementary material Fig. S5).

We therefore decided to analyze the SANS curve from the sample containing equimolar amounts (270 μM) of βγ and KL/EC as representing the βγ(KL/EC) complex. The scattering curve $I(Q)$ has a rather featureless profile (Fig. 4A) typical of elongated particles. The radius of gyration R_g was found to be 47±3 Å. We have also determined the cross-sectional radius of gyration R_C to be 9.7±0.6 Å (supplementary material Fig. S6). The profile of the distance distribution function $P(r)$ is similar to those produced by rod-like structures (45), with a steep increase at short distances and a slow decay at larger distances (Fig. 4B). The optimal D_{max} was found to be 160 Å, which is consistent with the R_g and R_C values (equation 6, Experimental procedures). Five independent low resolution models of the βγ(KL/EC) complex were calculated *ab initio* from the SANS curve (Figs 4C and supplementary material S7). All the models are very elongated and straight with a small cross section diameter. A common feature found in all models is an excess of mass near one end of the elongated structure. The long dimension of the excess density is about 60 Å. It is tempting to interpret this bulge as the βγ-protein sitting near the tip of the (KL/EC) coiled-coil dimer.

The SAXS curve of a sample containing equimolar quantities of βγ and KL/EC (144 μM) was very similar to the SANS curve (Fig. 4A). Because of short exposure time due to radiation damage, the quality of the SAXS data was not significantly better than that of the SANS experiments.

We also carried out rigid-body modeling of the complex. Rigid-body modeling seeks to minimize iteratively the difference between the experimental curve and the theoretically calculated one of the model by translating and rotating the two partners as rigid bodies in a grid search. As rigid bodies, we used the low-resolution model of the βγ-fragment based on the SAXS data, and a high-resolution model of the dimerization domain of the cortexillin I from *Dictyostelium discoideum* as a proxy for the (KL/EC) coiled-coil dimer. The parallel homo-dimer coiled-coil of cortexillin comprises 200 amino acids (14 heptads) that form a long rod, which we used as a model of (KL/EC) that totals 216 amino acids and is predicted to contain at least 10 heptads plus N- and C-terminal extensions (the purification tags and the regions containing the few conserved amino acids, respectively). Fig. 4D shows five different models that fit the experimental SANS scattering curve equally well. In all models, the βγ-protein is positioned near the end of the elongated partner, their long axis being roughly aligned. The rigid body modeling did not allow to discriminate between the two alternative orientations (head up

or head down) of the βγ-fragment relative to the (KL/EC) rod.

Manual sampling of alternative positions of the βγ-protein (*e. g.* at the extreme tip or in the middle of the rod) and orientations (*e. g.* perpendicular to the rod) yielded theoretical curves with worse fit to the data (supplementary material Fig. S8).

Interaction between the extracellular part of DivIB and KL/EC investigated by surface plasmon resonance- To evaluate the affinity of DivIB for the dimer (FtsL/DivIC) we used surface plasmon resonance (Biacore™ technology), where one partner (the ligand) is attached to a sensor chip while the other interacting partner (the analyte) is presented over the sensor chip in a flowing solution. The surface plasmon resonance being sensitive to the mass of the particles attached to the sensor's surface, the association kinetics can then be measured. The dissociation kinetics can be measured afterwards by flowing a solution without analyte.

For the surface plasmon resonance experiments, the (KL/EC) complex was attached to the sensor chip by the N-termini in order to present the FtsL/DivIC moieties to the solution. Attachment was achieved on Ni^{++}-NTA chips with both KL and EC proteins harboring an N-terminal poly-His-tag. The KL/EC mixture was injected onto the sensor chip at a concentration were the (KL/EC) complex is a dimer, and not a tetramer. Association and dissociation curves could be recorded for the three soluble forms $DivIB_{ext}$, β and βγ.

The β and βγ truncated forms produced comparable interaction curves with (KL/EC) (not shown). We measured the kinetic parameters in the case of the βγ-protein. Data are shown in Fig. 5. The theoretical curves describing the interaction according to a simple Langmuir model of association, βγ + (KL/EC) ⇆ βγ(KL/EC), were globally fitted to the experimental curves. The extracted association rate constant was k_A = 7.3 10^4 $M^{-1}s^{-1}$, the dissociation rate constant k_D = 0.016 s^{-1} and the equilibrium dissociation constant was K_D = 220 nM, with χ^2 = 2.26.

In the case of $DivIB_{ext}$, the association curve did not reach a plateau, but continued to increase linearly, indicating aggregation (not shown). A similar linear increase was observed when $DivIB_{ext}$ was injected on the sensor chip devoid of (KL/EC). However, the aggregation rates on the naked surface and on the (KL/EC)-coated surface were not identical in most experiments, thus precluding subtraction. In one instance, when the linear increase was identical in the presence and the absence of (KL/EC), an equilibrium constant of about 300 nM was determined, which is comparable to the constant measured with the βγ-protein.

The value of the affinity of the βγ-fragment for the (KL/EC) dimer has no direct significance for the physiological interaction between DivIB and (FtsL/DivIC), as the proteins *in vivo* are attached to the plasma membrane and thus restricted in two-dimensions. However, the relatively high dissociation constant explains the practical difficulty in isolating the recombinant complex *in vitro*, as it can be partially dissociated during washing steps of co-purification procedures, or during the course of a size exclusion chromatography.

To examine the importance of the C-terminal regions of FtsL and DivIC in the interaction with DivIB, we produced truncated variants of KL and EC, termed KL* and EC*, and tested the interaction of the KL/EC* and KL*/EC dimer with the βγ-fragment using the surface plasmon resonance assay. The KL and EC protein were truncated so as to preserve the predicted coiled-coil region (Fig. 6A).

Fig. 6B shows the surface plasmon resonance interaction curves of the βγ-fragment with immobilized (KL/EC), (KL*/EC) and (KL/EC*). Care was taken to immobilize comparable amounts of each dimer onto the sensor chip. Strikingly, truncation of either KL or EC abolished completely the interaction with the βγ-protein, demonstrating the importance of the C-terminal region of both FtsL and DivIC.

DISCUSSION

DivIB, FtsL and DivIC form a ternary complex with an enigmatic function that is central to the process of bacterial division. The present work provides the first structural information on the organization of the complex, suggesting new hypotheses.

The crystal structures of the POTRA and β domains from *E. coli* and *Y. enterolytica* FtsQ have been solved (24). In contrast, our NMR and proteolysis results show that the POTRA domain from *S. pneumoniae* is not folded in the recombinant soluble form of the protein, as was observed with DivIB from *G. stearothermophilus* (26). We find it unlikely that the DivIB POTRA domain from the two latter species is truly unfolded under physiological conditions, given the conservation of the POTRA motif (25) and the results from sequence hydrophobic cluster analysis that are typical of folded proteins (11). It is more likely that the DivIB POTRA domain has a marginal conformational stability in some organisms, and that this domain is not stable in conditions that differ from those *in vivo*. Factors that could stabilize the structure of the POTRA domain *in vivo* include the pH, which may be different close to the membrane than in the bulk water, the ionic strength, the charges at the membrane surface, or the presence of interacting protein partners. Indeed, epitopes required for the recruitment of DivIB(FtsQ) to the division site have been found in the POTRA domain, suggesting interactions with other proteins (23,24).

The low-resolution structure of the β-domain of pneumococcal DivIB resembles more that from *E. coli* FtsQ than that from *G. stearothermophilus* DivIB. The main reason is that the β-domain from *G. stearothermophilus* is shorter as it lacks two C-terminal β-strands (strands S11 and S12, Fig. 7A). Our proteolysis and NMR data concur with the crystallographic results to conclude that the β-domain includes S11 and S12. The absence of these β-strands in the β-domain from *G. stearothermophilus* is likely the unfortunate consequence of the accessible loop between S11 to the helix H5 containing a trypsin cleavage site, as proposed previously (24).

The NMR spectra of $DivIB_{ext}$ with and without (KL/EC) showed that the β-domain is involved in the interaction. The signal from the unfolded α- and γ-regions were not significantly affected, implying little or no interaction of these parts with the (KL/EC) dimer. However, as proton resonances of the unfolded segments are not resolved, the disappearance of some peaks from this region of the spectrum would pass unnoticed. We therefore cannot rule out completely that formation of the $DivIB_{ext}$(KL/EC) complex induces some marginal structure in the α- or γ-regions.

Using a deuterated β-domain to overcome relaxation problems resulting from the hydrodynamic properties of the β(KL/EC) complex, it has been possible to detect the modification of chemical shift of 27 amide protons, which could be identified following attribution of the resonances. Given the good agreement between the low-resolution structure of the β-domain from pneumococcal DivIB and the crystal structure from *E. coli* and *Y. enterocolitica* FtsQ, we have tried to locate the positions affected by the interaction with (KL/EC). For that purpose, we constructed an atomic model of the β-domain from *S. pneumoniae* based on that from *Y. enterocolitica* using the program SWISS-MODEL (46). As the sequence identity of 10% between the β-domains of both organisms is very low, the proposed model should be considered with caution and certainly not as a high-resolution structure. However, the hypothetical model presented in Figs 7A and 7B is likely sufficient to broadly locate the regions of interaction with FtsL and DivIC. Indeed, the sequence alignment between the *S. pneumoniae* and *Y. enterolytica* proteins resulting from the model building, and the classical alignment between the *S. pneumoniae* and *G. stearothermophilus* show a good agreement for the position of the secondary structure elements (Fig 7C), lending some credence to the model.

In the model of the *S. pneumoniae* β-domain, the positions affected by the interaction with (KL/EC), are defining a large surface on the face of the "bean" that is formed by the β-sheet, whereas the opposite face formed by the α-helices is largely unaffected (Figs 7A and B). Previous mutagenesis studies *in vivo* have implicated the C-terminal β-strands of the β-domain in the recruitment of (FtsL/FtsB) in *E. coli* (24), and in the function of pneumococcal DivIB likely linked to the interaction with FtsL (17). Unfortunately, we were not able to attribute unambiguously the amide protons of this part of the sequence to NMR resonances (Fig. 2B). The absence of resonance peaks is likely due to conformational exchanges in this region of the protein. Residues 353-361 with attributed resonance are not included in the model due to the uncertainty of the alignment in the

region. It is therefore possible that the interface with the (KL/EC) dimer extends to the C-terminal region at the "top" of the "bean".

In the absence of high resolution structure of the pneumococcal β-domain, it is premature to test the importance of residues at the interface with FtsL and DivIC by site-directed mutagenesis. As the interface covers a large part of one face of the β-domain, it is possible that single mutations would not be sufficient to suppress the interaction, although the D237N and A252P substitutions in *E. coli* FtsQ, each abolished the recruitment of FtsL to the division site (24,47).

The *ab initio* model of the βγ(KL/EC) complex based on the SANS data, corroborated by the rigid body modeling, indicates that the β-domain is positioned near the tip of the rod constituted by the long (KL/EC) coiled-coil. Previous controls had shown that $DivIB_{ext}$ does not interact with the artificial K/E part (11). It was therefore reasonable to assume that the β-domain binds to the C-terminal regions of FtsL and DivIC. This result was confirmed by the complete loss of the interaction measured by surface plasmon resonance when either KL or EC were truncated of their C-termini.

When viewing the models presented in Fig. 4, it is to be remembered that a large part of the long rod is constituted by the K and E coiled-coil moieties that were added to constrain the dimerization of the extracellular regions of FtsL and DivIC (11). The purification tags of the KL and EC proteins are also present and contribute to the length of the complex. However, the five predicted heptads of the extracellular coiled-coil of the FtsL/DivIC dimer should extend about 50 Å from the membrane to present the C-terminal regions that contains the conserved residues. The POTRA or α-domain of *E. coli* FtsQ is about 30 Å long and there are about eight additional residues between the transmembrane segment and the POTRA domain, so that the lower part of the β-domain could stand about 30 to 60 Å from the membrane, whether the eight additional amino acids are tightly folded or fully extended. However, the sequence of the eight residues proximal to the membrane does not have the polar character to constitute a long solvent-exposed extended linker. If the base of the β-domain does not reach 50 Å from the membrane, the (FtsL/DivIC) rod may need to be tilted to present its extremity to the β-domain. An angle of 37 ° would allow the end of the FtsL/DivIC coiled-coil to make contact with the base of a β-domain 40 Å from the membrane. Note that tilt angles of this magnitude are not unheard of for transmembrane helices (48). Alternatively, the presence of a kink between the transmembrane segment and the extracellular coiled-coil of FtsL and DivIC cannot be excluded. Indeed, the prediction score for a helical structure, using various programs, falls in the segment of a few residues next to the membrane (Fig. 8).

Several genetic studies have failed to indicate a role of the POTRA domain in the recruitment of FtsL and DivIC(FtsB). Also, our NMR data do not support an interaction between the POTRA domain and the (FtsL/DivIC) dimer. If the C-terminal segments of the dimer constitute the sole region that interacts with the β-domain of DivIC(FtsB), the resulting model leaves the coiled-coil rod free for interacting with other proteins. In this respect, it is striking that the Campylobacterales, which divide without FtsQ, have retained the coiled-coil part of FtsL and FtsB. The absence of the C-terminal regions of FtsL and FtsB in Campylobacterales indicates that the coiled-coil region plays a role *per se*, and not simply as a way to bring the C-terminal parts together. The coiled-coil rod of the FtsL/DivIC(FtsB) dimer, for example, could interact with the septal penicillin-binding protein FtsI, to regulate the function of this enzyme in the assembly of the cell wall. Indeed, a dimer of FtsL and a C-terminally truncated form of FtsB is still able to recruit FtsI to the division site (27).

It is paradoxical that the coiled-coil part of the FtsL/DivIC(FtsB) dimer appears to be the essential feature of this complex as it exhibits no sequence conservation. This paradox prompts us to propose another speculative function for the dimer. One of the many processes coordinated during bacterial cell division is the membrane invagination. It is possible that the onset of the invagination, which implies a sharp membrane curvature, might require or benefit from a destabilization or a modification of the lipid bilayer properties. In this respect, it is noteworthy that the oligomerization of bitopic membrane proteins, driven by the formation of coiled-coils, is at the core of eukaryotic and viral membrane

fusion events (49). Most interestingly, it was found in biophysical studies that the oligomerization of a transmembrane fusogenic peptide, constrained by an artificial extra-membranous coiled-coil domain, modifies and destabilizes the membrane (50). An analogous phenomenon could occur during bacterial cell division, where the dimerization of FtsL and DivIC(FtsB) might bring in close proximity their transmembrane segments, thereby inducing a local modification of the membrane properties, which could be important for the invagination.

In these hypothetical models of the function of FtsL and DivIC, the capping provided by DivIB may simply stabilize the coiled-coil. This could explain why DivIB is essential only at high temperature in *B. subtilis* (8,19), if the (FtsL/DivIC) dimer of this species were stable enough at low temperature. This model could also explain why Campylobacterales can divide without FtsQ, if the interaction between their FtsL and FstB were strong enough. We propose to reverse our way of considering the interaction between FtsL and DivIC(FtsB). Instead of viewing the transmembrane segments and coiled-coil regions as means of bringing together the C-terminal parts, the interaction of the C-terminal parts of FtsL and DivIC(FtsB) with DivIB(FtsQ) could help to zip together the coiled-coil helices and bring the transmembrane segments together.

We have seen no structuration of the α-domain upon binding of (KL/EC), which suggests the absence of interaction of the POTRA domain with the (FtsL/DivIC) dimer. No mutation affecting the recruitment of FtsL/DivIC(FtsB) to the division site was found in the POTRA domain of *E. coli* FtsQ (24). Also, the complete POTRA domain can be deleted from *B. subtilis* DivIB (23) without compromising growth. In contrast, mutations in the POTRA domains have been found to affect the recruitment of FtsQ and DivIB to the division site of *E. coli* and *B. subtilis* (23,24). Could the POTRA domain intervene in the function of FtsL and DivIC(FtsB)? POTRA domain stands for polypeptide-transport-associated domain (25). Apart from DivIB(FtsQ), POTRA domains are found in β-barrel outer membrane proteins of Gram-negative bacteria, where they are thought to assist the folding of β-barrels by binding the unfolded polypeptide chain in an extended conformation (51). Coiled-coils, such as that found in FtsL and DivIC(FtsB), are assemblies of α-helices that are usually not folded individually prior to assembly. Thus, the extracellular region of FtsL and DivIC(FtsB) are likely to be unfolded without their partners (28). The POTRA domain could offer a binding interface, much as a chaperone, to FtsL and/or DivIC prior to their association. FtsL would be the preferred substrate of this chaperoning function of the POTRA domain, as it appears in two-hybrid experiments that FtsL alone interacts with FtsQ, whereas FtsB requires the presence of FtsL to interact with FtsQ (13). In these experiments, the interaction of FtsL might involve the POTRA domain, while the (FtsL/FtsB) dimer would interact with the β-domain of FtsQ.

The interaction of the β-domain of DivIB with the C-terminal regions of the FtsL/DivIC dimer revealed by our structural and interaction studies *in vitro*, opens new directions of investigation by suggesting that the coiled-coil region of FtsL and DivIC is free to interact with other partners, or plays a central role in itself.

REFERENCES

1. den Blaauwen, T., de Pedro, M. A., Nguyen-Disteche, M., and Ayala, J. A. (2008) *FEMS Microbiol Rev* **32**, 321-344
2. Harry, E., Monahan, L., and Thompson, L. (2006) *Int Rev Cytol* **253**, 27-94
3. Errington, J., Daniel, R. A., and Scheffers, D. J. (2003) *Microbiol Mol Biol Rev* **67**, 52-65
4. Goehring, N. W., and Beckwith, J. (2005) *Curr Biol* **15**, R514-526
5. Aarsman, M. E., Piette, A., Fraipont, C., Vinkenvleugel, T. M., Nguyen-Disteche, M., and den Blaauwen, T. (2005) *Mol Microbiol* **55**, 1631-1645
6. Buddelmeijer, N., Judson, N., Boyd, D., Mekalanos, J. J., and Beckwith, J. (2002) *Proc Natl Acad Sci USA* **99**, 6316-6321
7. Ghigo, J. M., Weiss, D. S., Chen, J. C., Yarrow, J. C., and Beckwith, J. (1999) *Mol Microbiol* **31**, 725-737
8. Daniel, R. A., and Errington, J. (2000) *Mol Microbiol* **36**, 278-289
9. Katis, V. L., Wake, R. G., and Harry, E. J. (2000) *J Bacteriol* **182**, 3607-3611
10. Buddelmeijer, N., and Beckwith, J. (2004) *Mol Microbiol* **52**, 1315-1327
11. Noirclerc-Savoye, M., Le Gouellec, A., Morlot, C., Dideberg, O., Vernet, T., and Zapun, A. (2005) *Mol Microbiol* **55**, 413-424
12. Daniel, R. A., Noirot-Gros, M. F., Noirot, P., and Errington, J. (2006) *J Bacteriol* **188**, 7396-7404
13. Karimova, G., Dautin, N., and Ladant, D. (2005) *J Bacteriol* **187**, 2233-2243
14. Daniel, R. A., Williams, A. M., and Errington, J. (1996) *J Bacteriol* **178**, 2343-2350
15. Guzman, L. M., Barondess, J. J., and Beckwith, J. (1992) *J Bacteriol* **174**, 7716-7728
16. Levin, P. A., and Losick, R. (1994) *J Bacteriol* **176**, 1451-1459
17. Le Gouellec, A., Roux, L., Fadda, D., Massidda, O., Vernet, T., and Zapun, A. (2008) *J Bacteriol* **190**, 4501-4511
18. Carson, M. J., Barondess, J., and Beckwith, J. (1991) *J Bacteriol* **173**, 2187-2195
19. Beall, B., and Lutkenhaus, J. (1989) *J Bacteriol* **171**, 6821-6834
20. Guzman, L. M., Weiss, D. S., and Beckwith, J. (1997) *J Bacteriol* **179**, 5094-5103
21. Katis, V. L., and Wake, R. G. (1999) *J Bacteriol* **181**, 2710-2718
22. Scheffers, D. J., Robichon, C., Haan, G. J., den Blaauwen, T., Koningstein, G., van Bloois, E., Beckwith, J., and Luirink, J. (2007) *J Bacteriol* **189**, 7273-7280
23. Wadsworth, K. D., Rowland, S. L., Harry, E. J., and King, G. F. (2008) *Mol Microbiol* **67**, 1143-1155
24. van den Ent, F., Vinkenvleugel, T. M., Ind, A., West, P., Veprintsev, D., Nanninga, N., den Blaauwen, T., and Lowe, J. (2008) *Mol Microbiol* **68**, 110-123
25. Sanchez-Pulido, L., Devos, D., Genevrois, S., Vicente, M., and Valencia, A. (2003) *Trends Biochem Sci* **28**, 523-526
26. Robson, S. A., and King, G. F. (2006) *Procs Natl Acad Sci USA* **103**, 6700-6705
27. Gonzalez, M. D., and Beckwith, J. (2009) *J Bacteriol* **10.1122B/JB.01597-08**
28. Robson, S. A., Michie, K. A., Mackay, J. P., Harry, E., and King, G. F. (2002) *Mol Microbiol* **44**, 663-674
29. Chao, H., Houston, M. E., Jr., Grothe, S., Kay, C. M., O'Connor-McCourt, M., Irvin, R. T., and Hodges, R. S. (1996) *Biochemistry* **35**, 12175-12185
30. Jung, Y. S., and Zweckstetter, M. (2004) *J Biomol NMR* **30**, 11-23
31. ESRF. (2009) ID2 High Brilliance Beamline. http://www.esrf.eu/UsersAndScience/Experiments/SCMatter/ID02/.
32. Hammersley, A. (2004) The Fit2D home page. http://www.esrf.eu/computing/scientific/FIT2D/.
33. ILL. (2008) Large dynamic range small-angle diffractometer D22. http://www.ill.eu/d22/ .
34. Ghosh, R., Egelhaaf, S. U., and Rennie, A. R. (2006) A computing guide for small-angle scattering experiments. ILL Report No. ILL06GH05T. Institut Laue Langevin, Grenoble.

35. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) *J Appl Cryst* **36**, 1277-1282
36. Mylonas, E., and Svergun, D. I. (2007) *J Appl Cryst* **40**, s245-s249
37. Jacrot, B., and Zaccai, G. (1981) *Biopolymers* **20**, 2413–2426
38. Guinier, A. (1939) *Ann Phys* **12**, 166–237
39. Semenyuk, A. V., and Svergun, D. I. (1991) *J Appl Cryst* **24**, 537–540
40. Svergun, D. I. (1999) *Biophys J* **76**(6), 2879-2886
41. Konarev, P. V., Petoukhov, M. V., and Svergun, D. I. (2001) *J Appl Cryst* **34**, 527-532
42. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J Mol Graphics* **14**, 51-55
43. Svergun, D. I., Richard, S., Koch, M. H., Sayers, Z., Kuprin, S., and Zaccai, G. (1998) *Proc Natl Acad Sci USA* **95**, 2267-2272
44. Svergun, D. I., Barberato, C., and Koch, M. H. J. (1995) *J Appl Cryst* **28**, 768-773
45. Glatter, O. (1979) *J Appl Cryst* **12**, 166-175
46. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) *Bioinformatics* **22**, 195-201
47. Goehring, N. W., Petrovska, I., Boyd, D., and Beckwith, J. (2007) *J Bacteriol* **189**, 633-645
48. Nyholm, T. K., Ozdirekcan, S., and Killian, J. A. (2007) *Biochemistry* **46**, 1457-1465
49. Harrison, S. C. (2008) *Nat Struct Mol Biol* **15**, 690-698
50. Lau, W. L., Ege, D. S., Lear, J. D., Hammer, D. A., and DeGrado, W. F. (2004) *Biophys J* **86**, 272-284
51. Kim, S., Malinverni, J. C., Sliz, P., Silhavy, T. J., Harrison, S. C., and Kahne, D. (2007) *Science* **317**, 961-964
52. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res* **22**, 4673-4680

FOOTNOTES

We thank Jean-Pierre Andrieux for the N-terminal sequencing and Bernard Dublet for mass spectrometry analysis. We are grateful to Isabelle Petit-Härtlein, Nicole Thielens and Christine Ebel for assistance with protein production, surface plasmon resonance and AUC experiments, respectively, and to Anuj Shukla, Theyencheri Narayanan and Stéphanie Finet for their help with the SAXS experiments. This work was supported by the European grant LSMH-CT-EUR-INTAFAR 2004-512138 to Vernet's and Simorre's laboratories, by the Ministry of Research "allocation de recherche" to S. Masson, and by EPSRC grants GR/R99393/01 and EP/C015452/1 to P. Callow.

FIGURE LEGENDS

Fig. 1. Trypsin digestion of the extracellular region of pneumococcal DivIB. $DivIB_{ext}$ was incubated with trypsin at a 100:1 (w/w) ratio at 37 °C. Aliquots withdrawn after the indicated time intervals (in min) were analyzed by Coomassie-stained SDS polyacrylamide gel electrophoresis. Molecular weights in kDa are given on the left. After 75 min of incubation, a fragment migrating as a 15 kDa protein remains.

Fig. 2. Interaction of the β-domain with (KL/EC) monitored by NMR. A. HSQC [$^{15}N;^{1}H$] of $^{15}N/^{2}H$-β-domain without (black) and with (KL/EC). B. Sequence of the pneumococcal β-domain. Residues with attributed amide proton resonances are colored. Residues with chemical shift modification of their amide proton are in red.

Fig. 3. Small angle X-ray scattering study of the β- and βγ-domains of DivIB. A. Experimental scattering curve of pneumococcal β in blue and βγ in red, and theoretical curves of the β-domains from *E. coli* in green and *G. stearothermophilus* in gray. $I(0)$ values were normalized to 1. The noise in the data

at higher *Q*-values is due to the short exposure time used to limit radiation damage. B. Distance distribution functions *P*(*r*) corresponding to the scattering curves shown in A. C. Envelope of the high-resolution structures of the β-domains from *G. stearothermophilus* in gray and *E. coli* in green (from the NMR structure PDB # 1YR1, and X-ray structure PDB # 2VH1, respectively), and low-resolution models calculated ab initio from the distance distribution functions of the pneumococcal β-domain (in blue) and βγ-fragment (in red). D. Superimposition of the envelope of the low-resolution SAXS model of the βγ-fragment from *S. pneumoniae* and the ribbon representation of the β-domain from *E. coli*.

Fig. 4. Small-angle neutron and x-ray scattering study of the βγ(KL/EC) complex. A. SANS (red) and SAXS (black) curves *I*(*Q*). B. Distance distribution function *P*(*r*) from SANS data. C. Three dimensional model of the βγ(KL/EC) complex calculated *ab initio* from the SANS distance distribution function. The four views correspond 90° rotation steps around the long axis. Note that the spheres do not represent atoms or amino acids, but are dummy scattering particles. D. Rigid body modeling of the βγ(KL/EC) complex using the low-resolution model of the βγ-protein obtained by SAXS (in red), and the crystal structure of cortexilin 1 (in blue) as a model for the (KL/EC) dimer. Five models calculated by fitting their theoretical curve to the experimental data. Fits are shown below their respective models.

Fig. 5. Interaction between (KL/EC) and the βγ-fragment monitored by surface plasmon resonance. The (KL/EC) dimer (400 RU) was attached by its poly-His-tags to a Ni-NTA sensor chip. The βγ-protein was injected at the concentrations of 0.08 mM, 0.15 mM, 0.30 mM, 0.55 mM, 0.77 mM, 1.00 mM and 1.46 mM, corresponding to the curves from lowest to highest. The association and dissociation rate constant were extracted by globally fitting theoretical curves to the data according to a simple Langmuir model of association. Thin lines show the theoretical fitted curves.

Fig. 6. Interaction between the (KL/EC) dimer, or the truncated forms (KL*/EC) and (KL/EC*), and the βγ-protein monitored by surface plasmon resonance. A. Linear representation of the three dimers investigated, comprising the purification tags, the e and k artificial coiled-coils, and the full length or truncated extracellular region of FtsL (FL or FL*) and DivIC (DC and DC*). B. Sensorgram showing the attachment of equivalent amounts of the (KL/EC) (solid line), (KL*/EC) (dashed line) or (KL/EC*) (dotted line) on the left scale, and the association and dissociation of the βγ-fragment on the respective dimers on the right scale.

Fig. 7. A. Model of the envelope of the β-domain from *S. pneumoniae* produced with SWISS-MODEL using the crystal structure of *Y. enterocolitica* FtsQ (PDB #2VH2), the α-domain would be at the bottom, whereas the C-terminal γ-tail would be at the top of the image. Residues are colored according to corresponding positions in the pneumococcal protein that were investigated by NMR. Modification of chemical shift of the amide protons upon binding of (KL/EC) is noted in red. No modification of chemical shift is in blue, and the absence of assignement is in gray. B. Ribbon representation of the model as in A. C. Sequence aligment of the extracellular region of DivIB(FtsQ). The α-domain and γ-tail are in lower case, the β-domain is in upper case. Elements of secondary structure are indicated in yellow for β-strands and reddish for α-helices. The lighter colors for the pneumococcal sequence indicate the hypothetical nature of the structural model. The nomenclature for the secondary structure is that from *E. coli* (24). Sequences were aligned using CLUSTALW (52) including sequences from additional species, in particular for the α-domain, and structure superimposition for the β-domain (PDB #1YR1, #2VH1 and #2VH2). Positions in the pneumococcal β-domain that have modified chemical shift of their amide proton upon binding of (KL/EC) are in bold.

Fig. 8. Model of a membrane-bound ternary complex of DivIB(FtsQ), FtsL and DivIC(FtsB). The envelope of the α- and β-domains of DivIB(FtsQ) is from the crystal structure of the *E. coli* protein (PDB

#2VH1). The β-domain is in red, FtsL and DivIC(FtsB) are in blue and green. The length of the extracellular coiled-coil spans five heptads. The C-terminal regions of FtsL and DivIC(FtsB) are positioned against one face of the β-domain. The coiled-coil rod is free to interact with other partners, such as FtsI. The transmembrane segments of FtsL and DivIC(FtsB) may be brought in proximity by the formation of the coiled-coil to provide a local signal.

Table 1. Forward scattering values and associated apparent masses and radii of gyration from SANS measurements of KL/EC, βγ and KL/EC/βγ samples.

[βγ] (μM)	0	270	340
[KL/EC] (μM)	270	270	0
$I(0)$	0.201±0.010	0.305±0.010	0.085±0.009
M_{app} (kDa)	55±8	46±6	22±5

The error on mass includes an estimated 10% error on the protein concentration.

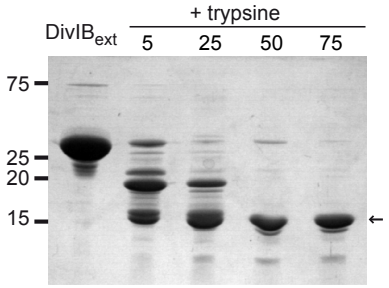
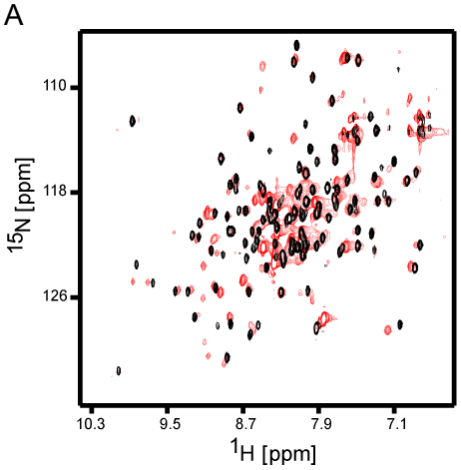


Fig. 1

A



B

```
224 DIVAYYISGE NHYPILSSGQ LETSSVSLNS
254 LPETYLSVLF NDSEQIKVFV SELAQISPEL
284 KAAIQKVELA PSKVTSDLIR LTMNDSDEVL
314 VPLSEMSKKL PYYSKIKPQL SEPSVVDMEA
344 GIYSYTVADK LIMEAEEK
```

Fig. 2

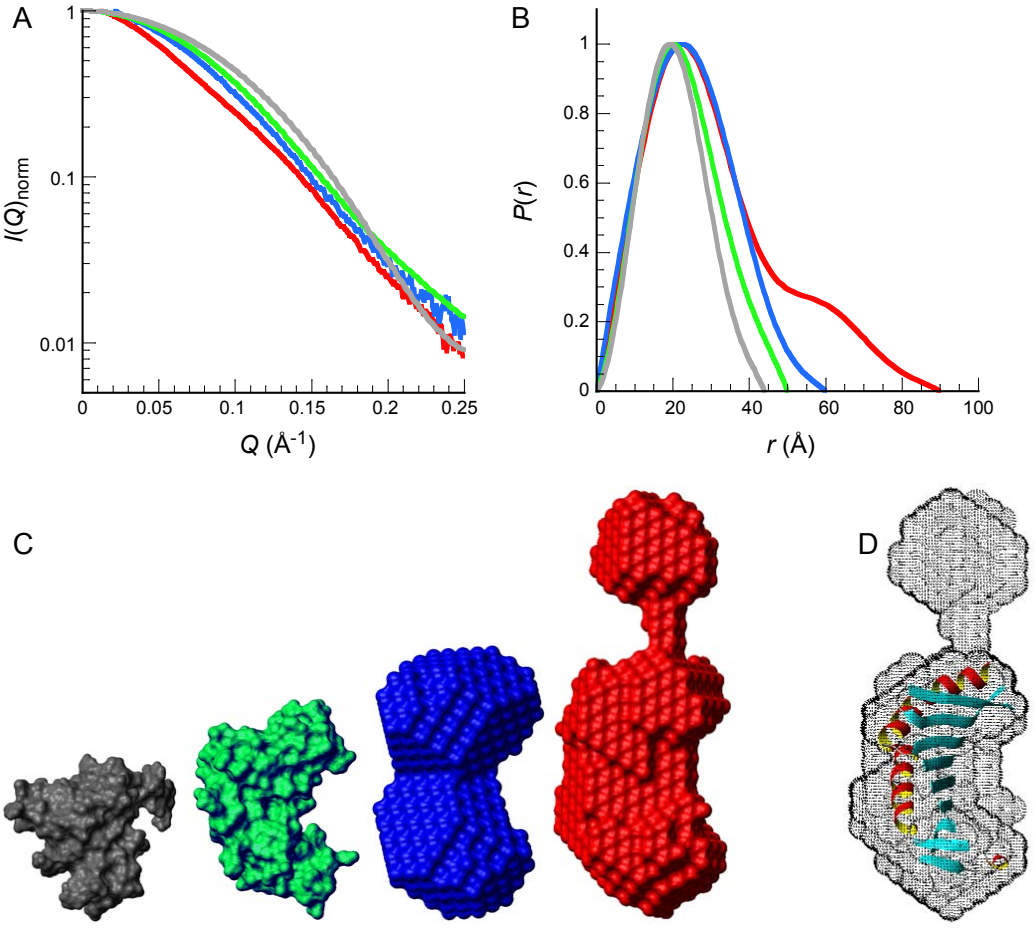


Fig. 3

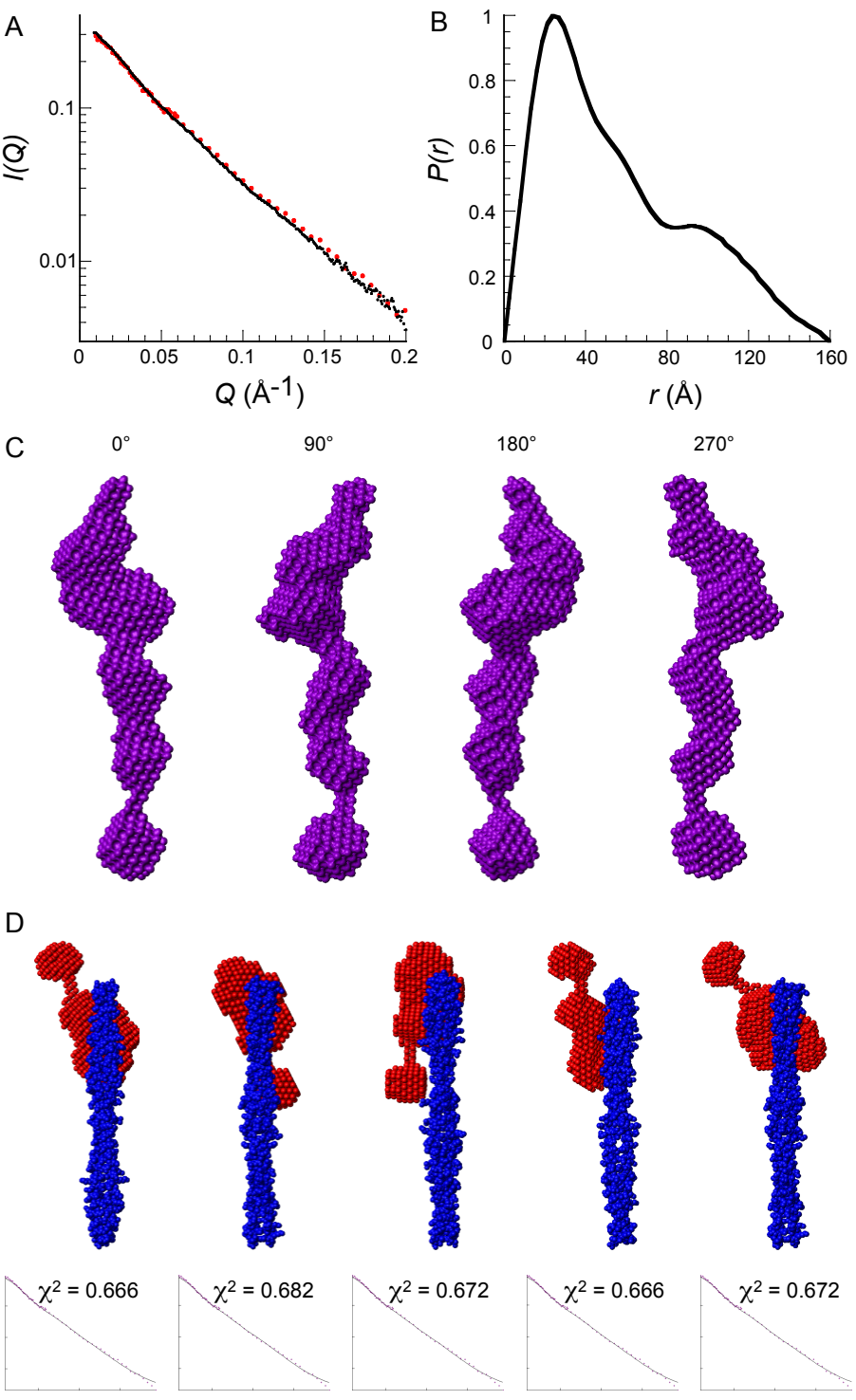


Fig. 4

Fig. 5

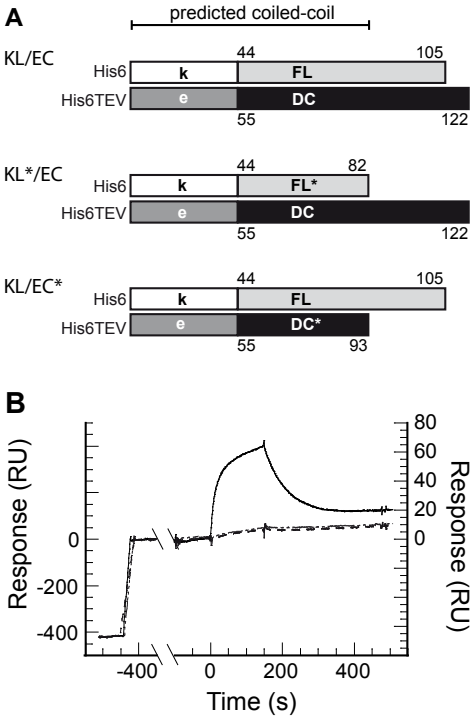
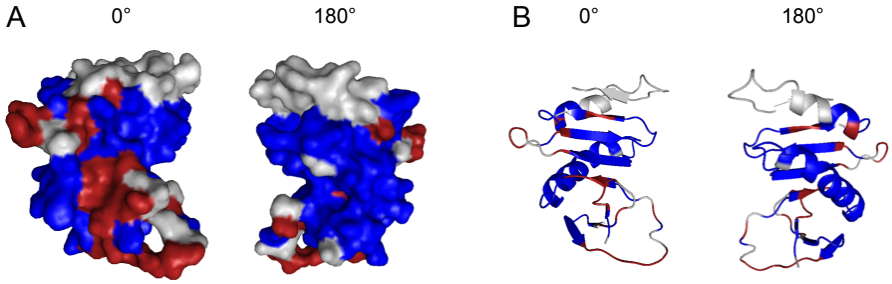


Fig. 6



C

```
G. stearotherm.  057 splgavghve vsgnrhltae riislsgitk r--tsfwkvn eqnvekkltr -heeikeatv ekqlpntiai
S. pneumoniae    150 spyatmkdir vegtvqttad dirqasgiqd sdytinllld kakyekqiks -nywvesaql vyqfptkfti
Y. enterocolitica 053 asrlplsklv vtgerhyttn ddirqailsl gapgtfmtqd vniiqqqier -lpwiqqasv rkqwpdelki
E. coli          052 aqrlplsklv ltgerhytrn ddirqsilal gepgtfmtqd vniiqtqieq rlpwikqvsv rkqwpdelki

G. stearotherm.  114 hvrewrRIAY VYDRQTFFPL LENGRLLKQE ETKTAPSDAP VLVGWKDGD- AIAEMTGQLA ELPAAVLGAM
S. pneumoniae    219 kVKEYDIVAY YISGENHYPI LSSGQLETSS VSLNSLPETY LSVLFNDSE- QIKVFVSELA QISPELKAAI
Y. enterocolitica 122 hlveYVPFAR WNDL---HMV DEQGRSFSVP SERVGKQKLP LLYGPEGSEQ DVLEGYRAIN KVLAANKYQL
E. coli          122 hlveYVPIAR WND---QHMV DAEGNTFSVP PERTSKQVLP MLYGPEGSAN EVLQGYREMG QMLAKDRFTL
                             S4        S5      S6              S7          H4

G. stearotherm.  183 SEIHYKPTRE YEDRVIVYMN DGYEVSATIR QFADKLSHYP AIAAALDRNV Kgvihlevgs yfvpysppkk
S. pneumoniae    288 QKVELAPSKV TSDLIRLTMN DSDEVLVPLS EMSKKLPYYS KIKPQLSEPS VVDMEAGIYS YTVADKLIME
Y. enterocolitica 189 KMVAMS--AR H--SWQLALD NDVRLELGRD DRMGRLQRFI ELYPMLQQQP --D-KRVSYV DLRYETGAAI
E. coli          189 KEAAMT--AR R--SWQLTLN NDIKLNLGRG DTMKRLARFV ELYPVLQQQA QTDGKRISYV DLRYDSGAAV
                     S8          S9          S10             H5                  S11     S12

G. stearotherm.  253 edgdettsp
S. pneumoniae    369 AEEKakqeak eaekkqeeeq kkqeeesnrn qtnqrssrr
Y. enterocolitica 252 GWAPVFIgsq ggeppvngqq nsnpqqnqaq akqq
E. coli          255 GWAPLPpees tqqqnqaqae qq
```

Fig. 7

The Journal of Biological Chemistry
ASBMB
jbc

Fig. 8

LIST OF SCIENTIFIC PUBLICATIONS

Kern, T., Schanda, P., Brutscher, B. (2008) Sensitivity enhanced IPAP-SOFAST-HMQC for fast-pulsing 2D NMR with reduced radiofrequency load. *J Magn Reson.* 190(2):333-338

Kern, T., Hediger, S., Mueller, P., Giustini, C., Joris, B., Bougault, C., Vollmer, W., Simorre, JP. (2008) Toward the characterization of peptidoglycan structure and protein-peptidoglycan interactions by solid-state NMR spectroscopy. *J Am Chem Soc.* 130(17):5618-5619

Gal, M., Kern, T., Schanda, P., Frydman, L., Brutscher, B., (2009) An improved ultra-fast 2D NMR experiment: towards atom-resolved real-time studies of protein kinetics at multi-Hz rates. *J Biomol NMR.* 43(1):1-10

Masson, S., Kern, T., Le Gouellec, A., Giustini, C., Simorre, JP., Callow, P., Thierry, V., Gabel, F., Zapun, A., (2009) The central domain of DivIB caps the C-terminal regions of the FtsL/DivIC coiled-coil rod. *J Biol Chem.* in print

Kern, T., Giffard, M., Hediger, S., Amoroso, A., Giustini, C., Joris, B., Bougault, C., Vollmer, W., Simorre, JP., (2009) Structure and flexibility of bacterial cell walls by solid-state NMR - evidence for cooperative binding of metal ions. *J Mol Biol.* submitted

Lescop, E., Kern, T., Brutscher, B. Guidelines for the use of band-selective radiofrequency pulses in heteronuclear NMR: Example of longitudinal-relaxationenhanced BEST-type 1H-15N correlation experiments. *J Magn Reson.* submitted

BIBLIOGRAPHY

[1] Akke, M., **2002**. NMR methods for characterizing microsecond to millisecond dynamics in recognition and catalysis. *Curr Opin Struct Biol* 12:642–7. (Cited p. 75.)

[2] Andersson, P., Weigelt, J., and Otting, G., **1998**. Spin-state selection filters for the measurement of heteronuclear one-bond coupling constants. *J Biomol NMR* 12:435–441. (Cited p. 31 and 50.)

[3] Atrih, A., Bacher, G., Allmaier, G., Williamson, M. P., and Foster, S. J., **1999**. Analysis of peptidoglycan structure from vegetative cells of Bacillus subtilis 168 and role of PBP 5 in peptidoglycan maturation. *J Bacteriol* 181:3956–66. (Cited p. 89 and 134.)

[4] Aue, W., Bartholdi, E., and Ernst, R., **1976**. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *Journal of Chemical Physics* 64:2229–2246. (Cited p. 12 and 13.)

[5] Baddiley, J., Hancock, I. C., and Sherwood, P. M. A., **1973**. X-ray photoelectron studies of magnesium ions bound to the cell walls of Gram-positive bacteria. *Nature* 243:43–45. (Cited p. 99.)

[6] Batley, M., Redmond, J. W., and Wicken, A. J., **1987**. Nuclear magnetic resonance spectra of lipoteichoic acid. *Biochim Biophys Acta* 901:127–37. (Cited p. 145, 148 and 151.)

[7] Bermel, W., Bertini, I., Felli, I., Picciolo, M., and Pierattelli, R., **2006**. ^{13}C-detected protonless NMR spectroscopy of proteins in solution. *Prog Nucl Magn Reson Spectrosc* 48:25–45. (Cited p. 31 and 39.)

[8] Beveridge, T. J. and Murray, R. G., **1980**. Sites of metal deposition in the cell wall of Bacillus subtilis. *J Bacteriol* 141:876–87. (Cited p. 99.)

[9] Billeter, M., Wagner, G., and Wüthrich, K., **2008**. Solution NMR structure determination of proteins revisited. *J Biomol NMR* 42:155–8. (Cited p. 11 and 13.)

[10] Birdsell, D. C., Doyle, R. J., and Morgenstern, M., **1975**. Organization of teichoic acid in the cell wall of Bacillus subtilis. *J Bacteriol* 121:726–34. (Cited p. 134.)

[11] Blackledge, M., **2005**. Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings. *Prog Nucl Magn Reson Spectrosc* 46:23–61. (Cited p. 9, 27 and 107.)

[12] Boudet, J., Chouquet, A., Chabhoune, A., Guistini, C., Joris, B., Simorre, J.-P., and Bougault, C., **2007**. ^{1}H, ^{13}C and ^{15}N resonance assignments of YajG, an *Escherichia coli* protein of unknown structure and function. *Biomol NMR Assign* 1:89–91. (Cited p. 139 and 164.)

[13] Bougault, C., Feng, L., Glushka, J., Kupce, E., and Prestegard, J., **2004**. Quantitation of rapid proton-deuteron amide exchange using hadamard spectroscopy. *J Biomol NMR* 28:385–390. (Cited p. 57 and 59.)

[14] Caldarelli, S. and Emsley, L., **1998**. Intrinsic asymmetry in multidimensional solid-state NMR correlation spectra. *J Magn Reson* 130:233–237. (Cited p. 165.)

[15] Chaput, C. and Boneca, I. G., **2007**. Peptidoglycan detection by mammals and flies. *Microbes Infect* 9:637–647. (Cited p. 87.)

[16] Chen, L., Olsen, R. A., Elliott, D. W., Boettcher, J. M., Zhou, D. H., Rienstra, C. M., and Mueller, L. J., **2006**. Constant-time through-bond ^{13}C correlation spectroscopy for assigning protein resonances with solid-state NMR spectroscopy. *J Am Chem Soc* 128:9992–9993. (Cited p. 119, 121, 132 and 165.)

[17] de Jonge, B. L., Chang, Y. S., Gage, D., and Tomasz, A., **1992**. Peptidoglycan composition in heterogeneous TN551 mutants of a methicillin-resistant *Staphylococcus aureus* strain. *J Biol Chem* 267:11255–11259. (Cited p. 93.)

[18] deAzevedo, E. R., Hu, W.-G., Bonagamba, T. J., and Schmidt-Rohr, K., **1999**. Centerband-Only Detection of Exchange: Efficient Analysis of Dynamics in Solids by NMR. *J Am Chem Soc* 121:8411–8412. (Cited p. 123 and 126.)

[19] Dempsey, C., **2001**. Hydrogen exchange in peptides and proteins using NMR-spectroscopy. *J Biomol NMR* 39:135–170. (Cited p. 9 and 57.)

[20] Deschamps, M. and Campbell, I., **2006**. Cooling overall spin temperature: Protein NMR experiments optimized for longitudinal relaxation effects. *J Magn Reson* 178:206–211. (Cited p. 28.)

[21] Ding, K. and Gronenborn, A., **2002**. Sensitivity enhanced E.COSY-type HSQC experiments for accurate measurements of one-bond ^{15}N-^{1}H^{N} and ^{15}N-^{13}C′ and two-bond ^{13}C′-^{1}H^{N} residual dipolar couplings in proteins. *J Magn Reson* 158:173–177. (Cited p. 35.)

[22] Dmitriev, B. A., Ehlers, S., and Rietschel, E. T., **1999**. Layered Murein revisited: a fundamentally new concept of bacterial cell wall structure, biogenesis and function. *Med Microbiol Immunol* 187:173–181. (Cited p. 96.)

[23] Dmitriev, B. A., Toukach, F. V., Holst, O., Rietschel, E. T., and Ehlers, S., **2004**. Tertiary structure of Staphylococcus aureus cell wall murein. *J Bacteriol* 186:7141–7148.

[24] Dmitriev, B. A., Toukach, F. V., Schaper, K.-J., Holst, O., Rietschel, E. T., and Ehlers, S., **2003**. Tertiary structure of bacterial murein: the scaffold model. *J Bacteriol* 185:3458–3468. (Cited p. 87 and 96.)

[25] Duma, L., Hediger, S., Brutscher, B., Boeckmann, A., and Emsley, L., **2003**. Resolution enhancement in multidimensional solid-state NMR spectroscopy of proteins using spin-state selection. *J Am Chem Soc* 125:11816–11817. (Cited p. 31.)

[26] Eisenmesser, E. Z., Bosco, D. A., Akke, M., and Kern, D., **2002**. Enzyme dynamics during catalysis. *Science* 295:1520–3. (Cited p. 75.)

[27] Emsley, L. and Bodenhausen, G., **1990**. Gaussian Pulse Cascades-New Analytical Functions for Rectangular Selective Inversion and in-Phase Excitation in NMR. *Chem Phys Letts* 165:469–476. (Cited p. 62.)

[28] Emsley, L. and Bodenhausen, G., **1992**. Optimization of Shaped Selective Pulses for NMR Using Quaternion Description of Their Overall Propagators. *J Magn Reson* 97:135–148. (Cited p. 62, 66 and 67.)

[29] Ernst, R. and Anderson, W., **1966**. Application of Fourier transform spectroscopy to magnetic resonance. *Rev Sci Instrum* 37:93. (Cited p. 15, 16 and 19.)

[30] Ernst, R., Bodenhausen, G., and Wokaun, G., **1987**. Principles of Nuclear Magnetic Resonance in One and Two Dimensions. Oxford. (Cited p. 15, 16, 19 and 28.)

[31] Farjon, J., Boisbouvier, J., Schanda, P., Pardi, A., Simorre, J.-P., and Brutscher, B., **2009**. Longitudinal-relaxation-enhanced NMR experiments for the study of nucleic acids in solution. *J Am Chem Soc* 131:8571–7. (Cited p. 63, 70 and 73.)

[32] Felli, I. C. and Brutscher, B., **2009**. Recent advances in solution NMR: fast methods and heteronuclear direct detection. *Chemphyschem* 10:1356–68. (Cited p. 10, 14 and 15.)

[33] Fiala, R. and Sklenar, V., **2007**. ^{13}C-detected NMR experiments for measuring chemical shifts and coupling constants in nucleic acid bases. *J Biomol NMR* 39:153–163. (Cited p. 39.)

[34] Freeman, R. and Kupce, E., **2003**. New methods for fast multidimensional NMR. *J Biomol NMR* 27:101–13. (Cited p. 14, 15 and 23.)

[35] Freeman, R. and Kupce, E., **2003**. New methods for fast multidimensional NMR. *J Biomol NMR* 27:101–113. (Cited p. 41.)

[36] Frydman, L., Scherf, T., and Lupulescu, A., **2002**. The acquisition of multidimensional NMR spectra within a single scan. *Proc Natl Acad Sci U S A* 99:15858–15862. (Cited p. 15, 24 and 41.)

[37] Fung, B., Khitrin, A., and Ermolaev, K., **2000**. An improved broadband decoupling sequence for liquid crystals and solids. *J Magn Reson* 142:97–101. (Cited p. 166.)

[38] Gabel, F., Simon, B., Nilges, M., Petoukhov, M., Svergun, D., and Sattler, M., **2008**. A structure refinement protocol combining NMR residual dipolar couplings and small angle scattering restraints. *J Biomol NMR* 41:199–208. (Cited p. 105.)

[39] Gal, M., Kern, T., Schanda, P., Frydman, L., and Brutscher, B., **2009**. An improved ultrafast 2D NMR experiment: towards atom-resolved real-time studies of protein kinetics at multi-Hz rates. *J Biomol NMR* 43:1–10. (Cited p. 24.)

[40] Gal, M., Mishkovsky, M., and Frydman, L., **2006**. Real-time monitoring of chemical transformations by ultrafast 2D NMR spectroscopy. *J Am Chem Soc* 128:951–956. (Cited p. 15 and 48.)

[41] Gal, M., Schanda, P., Brutsche, B., and Frydman, L., **2007**. UltraSOFAST HMQC-NMR and the repetitive acquisition of 2D protein spectra at Hz rates. *J Am Chem Soc* 129:1372–1377. (Cited p. 15, 19, 24 and 46.)

[42] Geen, H. and Freeman, R., **1991**. Band-Selective Radiofrequency Pulses. *J Magn Reson* 93:93–141. (Cited p. 33, 49, 62, 65, 66, 67 and 76.)

[43] Girardin, S. E., Boneca, I. G., Carneiro, L. A. M., Antignac, A., Jéhanno, M., Viala, J., Tedin, K., Taha, M., Labigne, A., and Zähringer, U., **2003**. Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. *Science* 300:1584–1587. (Cited p. 98 and 161.)

[44] Giraud, N., Blackledge, M., Böckmann, A., and Emsley, L., **2007**. The influence of nitrogen-15 proton-driven spin diffusion on the measurement of nitrogen-15 longitudinal relaxation times. *J Magn Reson* 184:51–61. (Cited p. 113.)

[45] Giraud, N., Blackledge, M., Goldman, M., Böckmann, A., Lesage, A., Penin, F., and Emsley, L., **2005**. Quantitative analysis of backbone dynamics in a crystalline protein from nitrogen-15 spin-lattice relaxation. *J Am Chem Soc* 127:18190–201. (Cited p. 113 and 117.)

[46] Glauner, B. and Hoeltje, J.-V., **1990**. Growth pattern of the murein sacculus of *Escherichia coli*. *J Biol Chem* 265:18988–18996. (Cited p. 96.)

[47] Glauner, B., Hoeltje, J.-V., and Schwarz, U., **1988**. The composition of the murein of Escherichia coli. *J Biol Chem* 263:10088–95. (Cited p. 87, 88, 89, 96 and 97.)

[48] Gram, H. C., **1884**. Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten. *Fortschritte der Medizin* 2:185–189. (Cited p. 86.)

[49] Guan, R. and Mariuzza, R. A., **2007**. Peptidoglycan recognition proteins of the innate immune system. *Trends Microbiol* 15:127–134. (Cited p. 87, 98 and 141.)

[50] Gullion, T. and Schaefer, J., **1989**. Detection of weak heteronuclear dipolar coupling by rotational-echo double-resonance nuclear magnetic resonance. *Adv Magn Reson* 13:57–83. (Cited p. 123.)

[51] Hansen, D. F., Vallurupalli, P., and Kay, L. E., **2008**. An improved 15N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J Phys Chem B* 112:5898–5904. (Cited p. 9.)

[52] Hansen, D. F., Vallurupalli, P., and Kay, L. E., **2008**. Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *J Biomol NMR* 41:113–20. (Cited p. 77.)

[53] Hartmann, S. and Hahn, E., **1962**. Nuclear Double Resonance in the Rotating Frame. *Phys Rev* 128:2042–2053. (Cited p. 112.)

[54] Harz, H., Burgdorf, K., and Hoeltje, J.-V., **1990**. Isolation and separation of the glycan strands from murein of *Escherichia coli* by reversed-phase high performance liquid chromatography. *Analytical Biochemistry* 190:120–128. (Cited p. 88 and 97.)

[55] Hayashi, K., **1975**. A rapid determination of sodium dodecyl sulfate with methylene blue. *Analytical Biochemistry* 67:503–506. (Cited p. 162.)

[56] Hayhurst, E. J., Kailas, L., Hobbs, J. K., and Foster, S. J., **2008**. Cell wall peptidoglycan architecture in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 105:14603–4608. (Cited p. 88.)

[57] Hediger, S., Meier, B., and Ernst, R., **1995**. Adiabatic passage Hartmann-Hahn cross-polarization in NMR under magic-angle sample-spinning. *Chemical Physics Letters* 240:449–456. (Cited p. 113 and 166.)

[58] Heptinstall, S., Archibald, A. R., and Baddiley, J., **1970**. Teichoic acids and membrane function in bacteria. *Nature* 225:519–521. (Cited p. 153.)

[59] Hiller, S., Garces, R. G., Malia, T. J., Orekhov, V. Y., Colombini, M., and Wagner, G., **2008**. Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science* 321:1206–10. (Cited p. 105.)

[60] Horne, D. S. and Tomasz, A., **1993**. Possible role of a choline-containing teichoic acidin the maintenance of normal cell shape and physiology in *Streptococcus oralis*. *J Bacteriol* 175:1717–1722. (Cited p. 144 and 161.)

[61] Hughes, A. H., Stow, M., Hancock, I. C., and Baddiley, J., **1971**. Function of teichoic acids and effect of novobiocin on control of Mg^{2+} at the bacterial membrane. *Nat New Biol* 229:56–57. (Cited p. 99.)

[62] Hurst, A., Hughes, A. H., Duckworth, M., and Baddiley, J., **1975**. Loss of D-alanine during sublethal heating of *Staphylococcus aureus* S6 and magnesium binding during repair. *J Gen Microbiol* 89:277–284. (Cited p. 153.)

[63] Jacob, G. S., Schaefer, J., and Wilson, G. E. J., **1985**. Solid-state 13C and 15N nuclear magnetic resonance studies of alanine metabolism in Aerococcus viridans (Gaffkya homari). *J Biol Chem* 260:2777–81. (Cited p. 123.)

[64] Jarymowycz, V. and Stone, M., **2006**. Fast time scale dynamics of protein backbones: NMR relaxation methods, applications and functional consequences. *Chem Rev* 106:1624–1671. (Cited p. 9 and 27.)

[65] Jeener, J., **1971**. Lecture presented at Ampere International Summer School II. (Cited p. 12 and 13.)

[66] Jeener, J., Meier, B., Bachmann, P., and Ernst, R. R., **1979**. Investigation of exchange processes by two-dimensional NMR spectroscopy. *J Chem Phys* 71:4546–4553. (Cited p. 118 and 123.)

[67] Kay, L. E., **2005**. NMR studies of protein structure and dynamics. *J Magn Reson* 173:193–207. (Cited p. 75.)

[68] Kay, L. E., Keifer, P., and Saarinen, T., **1992**. Pure Absortption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *J Am Chem Soc* 114:10663–10665. (Cited p. 68.)

[69] Kelemen, M. V. and Rogers, H. J., **1969**. Three-dimensional molecular models of the mucopeptide from Staphylococcus aureus. *J Gen Microbiol* 57:13–4. (Cited p. 95.)

[70] Kelemen, M. V. and Rogers, H. J., **1971**. Three-dimensional molecular models of bacterial cell wall mucopeptides (peptidoglycans). *Proc Natl Acad Sci U S A* 68:992–6. (Cited p. 95.)

[71] Kern, T., Hediger, S., Muller, P., Giustini, C., Joris, B., Bougault, C., Vollmer, W., and Simorre, J.-P., **2008**. Toward the characterization of peptidoglycan structure and protein-peptidoglycan interactions by solid-state NMR spectroscopy. *J Am Chem Soc* 130:5618–5619. (Cited p. 55, 131, 134 and 137.)

[72] Kern, T., Schanda, P., and Brutscher, B., **2008**. Sensitivity-enhanced IPAP SOFAST-HMQC for fast pulsing 2D NMR with reduced radiofrequency load. *J Magn Reson* 190:333–338. (Cited p. 19, 48, 50 and 71.)

[73] Kim, S. and Szyperski, T., **2003**. GFT NMR, a new approach to rapidly obtain precise high-dimensional NMR spectral information. *J Am Chem Soc* 125:1385–93. (Cited p. 15 and 22.)

[74] Kim, S. J., Cegelski, L., Preobrazhenskaya, M., and Schaefer, J., **2006**. Structures of Staphylococcus aureus cell-wall complexes with vancomycin, eremomycin, and chloroeremomycin derivatives by 13C19F and 15N19F rotational-echo double resonance. *Biochemistry* 45:5235–50. (Cited p. 123, 125 and 131.)

[75] Kim, S. J., Cegelski, L., Steuber, D., Singh, M., Dietrich, E., Tanaka, K. S. E., Parr, T. R. J., Far, A. R., and Schaefer, J., **2008**. Oritavancin exhibits dual mode of action to inhibit cell-wall biosynthesis in [**?**]. *J Mol Biol* 377:281–293.

[76] Kim, S. J. and Schaefer, J., **2008**. Hydrophobic side-chain length determines activity and conformational heterogeneity of a vancomycin derivative bound to the cell wall of Staphylococcus aureus. *Biochemistry* 47:10155–61. (Cited p. 123 and 131.)

[77] Klein, R., Hartmann, S., Egge, H., Behr, T., and Fischer, W., **1996**. The aqueous solution structure of a lipoteichoic acid from *Streptococcus pneumoniae* strain R6 containing 2,4-diamino-2,4,6-trideoxy-galactose: evidence for conformational mobility of the galactopyranose ring. *Carbohydr Res* 281:79–98. (Cited p. 144 and 148.)

[78] Koch, A. L., **1998**. Orientation of the peptidoglycan chains in the sacculus of Escherichia coli. *Res Microbiol* 149:689–701. (Cited p. 87.)

[79] Kotnik, M., Anderluh, P. S., and Prezelj, A., **2007**. Development of novel inhibitors targeting intracellular steps of peptidoglycan biosynthesis. *Curr Pharm Des* 13:2283–2309. (Cited p. 87.)

[80] Kubler-Kielb, J., Coxon, B., and Schneerson, R., **2004**. Chemical structure, conjugation and cross-reactivity of *Bacillus pumilus* SH18 cell wall polysaccharide. *J Bacteriol* 186:6891–6901. (Cited p. 144.)

[81] Kupce, E., Boyd, J., and Campbell, I. D., **1995**. Short selective pulses for biochemical applications. *J Magn Reson B* 106:300–3. (Cited p. 62, 66 and 67.)

[82] Kupce, E. and Freeman, R., **1994**. Wide-Band Excitation with Polychomatic Pulses. *J Magn Reson* 108:268–273. (Cited p. 28, 33 and 49.)

[83] Kupce, E. and Freeman, R., **1996**. Optimized adiabatic pulses for wideband spin inversion. *J Magn Reson A* 118:299–303. (Cited p. 33, 49, 66, 67 and 68.)

[84] Kupce, E. and Freeman, R., **2003**. Projection-reconstruction of three-dimensional NMR spectra. *J Am Chem Soc* 125:13958–9. (Cited p. 22.)

[85] Kupce, E. and Freeman, R., **2004**. Projection-reconstruction technique for speeding up multidimensional NMR spectroscopy. *J Am Chem Soc* 126:6429–40. (Cited p. 15.)

[86] Kupce, E., Nishida, T., and Freeman, R., **2003**. Hadamard NMR spectroscopy. *Prog Nucl Magn Reson Spectrosc* 42:95–122. (Cited p. 59.)

[87] Labischinski, H., Barnickel, G., Bradaczek, H., and Giesbrecht, P., **1979**. On the secondary and tertiary structure of murein. Low and medium-angle X-ray evidence against chitin-based conformations of bacterial peptidoglycan. *Eur J Biochem* 95:147–55. (Cited p. 95.)

[88] Lambert, P., Hancock, I., and Baddiley, J., **1975**. Influence of alanyl ester residues on the binding of magnesium ions to teichoic acids. *Biochim J* 151:671–676. (Cited p. 151, 152 and 153.)

[89] Lambert, P., Hancock, I., and Baddiley, J., **1975**. The interaction of magnesium ions with teichoic acid. *Biochim J* 149:519–524. (Cited p. 151.)

[90] Lee, D., Voegeli, B., and Pervushin, K., **2005**. Detection of C′ C^{α} correlations in proteins using a new time- and sensetivity-optimal experiment. *J Biomol NMR* 31:273–278. (Cited p. 35.)

[91] Lerche, M., Meissner, F., A. ans Poulsen, and Sorenson, O., **1999**. Pulse sequences for measurements of one-bond ^{15}N-^{1}H coupling constants in the protein backbone. *J Magn Reson* 140:259–263. (Cited p. 35.)

[92] Lesage, A., Bardet, M., and Emsley, L., **1999**. Through-Bond Carbon-Carbon Connectivities in Disordered Solids by NMR. *J Am Chem Soc* 121:10987–10993. (Cited p. 118, 119, 132 and 165.)

[93] Lescop, E. and Brutscher, B., **2009**. Highly automated protein backbone resonance assignment within a few hours: the "BATCH" strategy and software package. *J Biomol NMR* 44:43–57. (Cited p. 78.)

[94] Lescop, E., Schanda, P., and Brutscher, B., **2007**. A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J Magn Reson* 187:163–169. (Cited p. 11, 15, 19, 61, 62, 63, 70 and 73.)

[95] Lescop, E., Schanda, P., Rasia, R., and Brutscher, B., **2007**. Automated spectral compression for fats multidimensional NMR and increased time resolution in real-time NMR spectroscopy. *J Am Chem Soc* 129:2756–2757. (Cited p. 15, 22 and 59.)

[96] Levitt, M. H., **2001**. *Spin Dynamics*. Wiley. (Cited p. 105.)

[97] Levitt, M. H. and Ernst, R. R., **1983**. Improvement of pulse performance in NMR coherence transfer experiments. A compensated INADEQUATE experiment. *Mol Phys* 50:1109–1124. (Cited p. 118 and 119.)

[98] Lewandowski, J. R., Sein, J., Blackledge, M., and Emsley, L., **2009**. Anisotropic Collective Motion Contributes to Nuclear Spin Relaxation in Crystalline Proteins. *J Am Chem Soc* . (Cited p. 113 and 117.)

[99] Loria, J. P., Rance, M., and Palmer, A. G. r., **1999**. A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J Am Chem Soc* 121:2331–2332. (Cited p. 9.)

[100] Loria, J. P., Rance, M., and Palmer, A. G. r., **1999**. A TROSY CPMG sequence for characterizing chemical exchange in large proteins. *J Biomol NMR* 15:151–155. (Cited p. 9, 76 and 77.)

[101] Losonczi, J., Kupce, E., Gray, G., Sandor, P., and Brutscher, B., **2006**. Aligning ultrafast NMR spectroscopy and cold probe technology to reveal the mysteries of protein dynamics. *American Laboratory* 38:23–25. (Cited p. 30.)

[102] Magnet, S., Arbeloa, A., Mainardi, J.-L., Hugonnet, J.-E., Fourgeaud, M., Dubost, L., Marie, A., Delfosse, V., Mayer, C., Rice, L. B., and Arthur, M., **2007**. Specificity of L,D-transpeptidases from gram-positive bacteria producing different peptidoglycan chemotypes. *J Biol Chem* 282:13151–9. (Cited p. 92.)

[103] Magnet, S., Dubost, L., Marie, A., Arthur, M., and Gutmann, L., **2008**. Identification of the L,D-transpeptidases for peptidoglycan cross-linking in Escherichia coli. *J Bacteriol* 190:4782–5. (Cited p. 98 and 156.)

[104] Mainardi, J.-L., Hugonnet, J.-E., Rusconi, F., Fourgeaud, M., Dubost, L., Moumi, A. N., Delfosse, V., Mayer, C., Gutmann, L., Rice, L. B., and Arthur, M., **2007**. Unexpected inhibition of peptidoglycan LD-transpeptidase from Enterococcus faecium by the beta-lactam imipenem. *J Biol Chem* 282:30414–22. (Cited p. 98 and 156.)

[105] Mainardi, J.-L., Morel, V., Fourgeaud, M., Cremniter, J., Blanot, D., Legrand, R., Frehel, C., Arthur, M., Van Heijenoort, J., and Gutmann, L., **2002**. Balance between two transpeptidation mechanisms determines the expression of beta-lactam resistance in Enterococcus faecium. *J Biol Chem* 277:35801–7. (Cited p. 92.)

[106] Malmodin, D. and Billeter, M., **2005**. Signal identification in NMR spectra with coupled evolution periods. *J Magn Reson* 176:47–53. (Cited p. 41.)

[107] Mandelshtam, Taylor, and Shaka, **1998**. Application of the filter diagonalization method to one- and two-dimensional NMR spectra. *J Magn Reson* 133:304–12. (Cited p. 21.)

[108] Marion, D., **2005**. Fast acquisition of NMR spectra using Fourier transform of non-equispaced data. *J Biomol NMR* 32:141–50. (Cited p. 15, 20 and 21.)

[109] Matias, V. R., Al-Amoudi, A., Dubochet, J., and Beveridge, T. J., **2003**. Cryo-transmission electron microscopy of frozen-hydrated sections of *Escherichia coli* and *Pseudomonas aeruginosa*. *J Bacteriol* 185:6112–6118. (Cited p. 88.)

[110] Matias, V. R. and Beveridge, T., **2005**. Cryo-electron microscopy reveals native polymeric cell wall structure in *Bacillus subtilis* 168 and the existence of a periplasmic space. *Mol Microbiol* 56:240–251. (Cited p. 88, 134 and 148.)

[111] Matias, V. R. and Beveridge, T. J., **2006**. Native cell wall organization shown by cryo-electron microscopy confirms the existence of periplasmic space in *Staphylococcus aureus*. *J Bacteriol* 188:1011–1021. (Cited p. 88.)

[112] McDermot, A. E., **2009**. Structure and dynamics of membrane proteins by magic angle spinning solid-state NMR. *Annu Rev Biophys* 38:385–403. (Cited p. 131.)

[113] McDermott, A. E., **2004**. Structural and dynamic studies of proteins by solid-state NMR spectroscopy: rapid movement forward. *Curr Opin Struct Biol* 14:554–61. (Cited p. 113 and 117.)

[114] McDonald, C., Inohara, N., and Nunez, G., **2005**. Peptidoglycan signaling in innate immunity and inflammatory disease. *J Biol Chem* 280:20177–20180. (Cited p. 87 and 141.)

[115] Mehring, M., **1983**. *Principles of High Resolution NMR in Solids*. Springer-Verlag. (Cited p. 105 and 111.)

[116] Meissner, A., Duus, J., and Sorenson, O., **1997**. Spin-state-selective excitation. Application for E.COSY-type measurements of J^{HH}. *J Magn Reson* 128:92–97. (Cited p. 50.)

[117] Meissner, A., Schulte-Herbruggen, T., Briand, J., and Sorenson, O., **1998**. Double spin-state-selective coherence transfer. Apllication for two-dimensional selection of multiplet components with long transverse relaxation times. *Mol Phys* 95:1137–1142. (Cited p. 50 and 70.)

[118] Meroueh, S. O., Bencze, K. Z., Hesek, D., Lee, M., Fisher, J. F., Stemmler, T. L., and Mobashery, S., **2006**. Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proc Natl Acad Sci U S A* 103:4404–9. (Cited p. 87, 95 and 96.)

[119] Metz, G., Wu, X. L., and Smith, S. O., **1993**. Ramped-Amplitude Cross-Polarization in Magic-Angle-Spinning NMR. *J Magn Reson A* 110:219–227. (Cited p. 113.)

[120] Mishkovsky, M. and Frydman, L., **2005**. Interlaced Fourier transformation of ultrafast 2D NMR data. *J Magn Reson* 173:344–350. (Cited p. 55.)

[121] Mitchell, P. and Moyle, J., **1956**. Osmotic function and structure in bacteria. *J Symp Soc Gen Microbiol* 6:150–180. (Cited p. 86.)

[122] Mittermaier, A. and Kay, L., **2006**. Review-new tools provide new insights in NMR studies of protein dynamics. *Science* 312:224–228. (Cited p. 9, 27 and 76.)

[123] Mobli, M., Stern, A. S., and Hoch, J. C., **2006**. Spectral reconstruction methods in fast NMR: reduced dimensionality, random sampling and maximum entropy. *J Magn Reson* 182:96–105. (Cited p. 15 and 20.)

[124] Neuhaus, F. C. and Baddiley, J., **2003**. A continuum of anionic charge: structures and functions of D-alanyl-teichoic acids in gram-positive bacteria. *Microbiol Mol Biol Rev* 67:686–723. (Cited p. 99 and 153.)

[125] Orekhov, V. Y., Ibraghimov, I. V., and Billeter, M., **2001**. MUNIN: a new approach to multi-dimensional NMR spectra interpretation. *J Biomol NMR* 20:49–60. (Cited p. 21.)

[126] Ottiger, M., Delaglio, F., and Bax, A., **1998**. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J Magn Reson* 131:373–378. (Cited p. 31 and 50.)

[127] Palmer, A. G., Cavanagh, J., Wright, P. E., and Rance, M., **1991**. Sensitivity Improvement in Proton-Detected 2-Dimensional Heteronuclear Correlation NMR-Spectroscopy. *J Magn Reson* 93:151–170. (Cited p. 34, 51, 62, 68 and 76.)

[128] Pan, Y., Shenouda, N. S., Wilson, G. E., and Schaefer, J., **1993**. Cross-links in cell walls of Bacillus subtilis by rotational-echo double-resonance 15N NMR. *J Biol Chem* 268:18692–5. (Cited p. 123.)

[129] Patti, G. J., Kim, S. J., and Schaefer, J., **2008**. Characterization of the peptidoglycan of vancomycin-susceptible Enterococcus faecium. *Biochemistry* 47:8378–85. (Cited p. 123.)

[130] Peersen, O. B., Wu, X. L., and Smith, S. O., **1993**. Variable-Amplitude Corss-Polarization NMR. *J Magn Reson A* 104:334–339. (Cited p. 113.)

[131] Peiser, L., Mukhopadhyay, S., and Gordon, S., **2002**. Scavenger receptors in innate immunity. *Curr Opin Immunol* 14:123–128. (Cited p. 87 and 99.)

[132] Pelupessy, P., **2003**. Adiabatic single scan two-dimensional NMR spectroscopy. *J Am Chem Soc* 125:12345–12350. (Cited p. 49.)

[133] Pervushin, K., Riek, R., Wider, G., and Wuethrich, K., **1997**. Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* 94:12366–12371. (Cited p. 50.)

[134] Pervushin, K., Riek, R., Wider, G., and Wüthrich, K., **1997**. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* 94:12366–71. (Cited p. 68.)

[135] Pervushin, K., Voegeli, B., and Eletsky, A., **2002**. Longitudinal H-1 relaxation optimization in TROSY NMR spectroscopy. *J Am Chem Soc* 124:12898–12902. (Cited p. 15, 16, 28 and 46.)

[136] Peschel, A. and Sahl, H., **2006**. The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nat Rev Micro* 4:529–536. (Cited p. 99 and 153.)

[137] Rice, C. V. and Wickham, J. R., **2005**. Heterogeneous binding of lipoteichoic acid to the surface of titanium dioxide as determined with ^{31}P solid-state NMR spectroscopy. *J Am Chem Soc* 127:856–857. (Cited p. 131 and 148.)

[138] Rijnaarts, H. H. M., Norde, W., Lyklema, J., and Zehnder, A. J. B., **1995**. The isoelectric point of bacteria as an indicator for the presence of cell-surface polymers that inhibit adhesion. *Colloids and Surfaces B-Biointerfaces* 4:191–197. (Cited p. 153.)

[139] Roberts, M. F., Jacobson, G. R., Scott, P. J., Mimura, C. S., and Stinson, M. W., **1985**. ^{31}P-NMR studies of the oral pathogen *Streptococcus mutans*: observation of lipoteichoic acid. *Biochim Biophys Acta* 845:242–248. (Cited p. 100 and 148.)

[140] Rose, R. K. and Hogg, S. D., **1995**. Competitive binding of calcium and magnesium to *Streptococcal* lipoteichoic acid. *Biochimica Et Biophysica Acta-General Subject* 1245:94–98. (Cited p. 151 and 152.)

[141] Rovnyak, D., Hoch, J. C., Stern, A. S., and Wagner, G., **2004**. Resolution and sensitivity of high field nuclear magnetic resonance spectroscopy. *J Biomol NMR* 30:1–10. (Cited p. 15.)

[142] Sakellariou, D., Brown, S. P., Lesage, A., Hediger, S., Bardet, M., Meriles, C. A., Pines, A., and Emsley, L., **2003**. High-resolution NMR correlation spectra of disordered solids. *J Am Chem Soc* 125:4376–4380. (Cited p. 132.)

[143] Schanda, P., **2009**. Fast-pulsing longitudinal relaxation optimized techniques: Enriching the toolbox of fast biomolecular NMR spectroscopy. *Prog Nucl Magn Reson Spectrosc* 55:238–265. (Cited p. 14 and 15.)

[144] Schanda, P. and Brutscher, B., **2005**. Very fast two-dimensional NMR spectroscopy for real time investigation of dynamic events in proteins on the time scale of seconds. *J Am Chem Soc* 127:8014–8015. (Cited p. 10, 11, 16, 28, 41, 49, 50 and 73.)

[145] Schanda, P., Brutscher, B., Konrat, R., and Tollinger, M., **2008**. Folding of the KIX domain: characterization of the equilibrium analog of a folding intermediate using 15N/13C relaxation dispersion and fast 1H/2H amide exchange NMR spectroscopy. *J Mol Biol* 380:726–41. (Cited p. 75.)

[146] Schanda, P., Forge, V., and Brutscher, B., **2006**. HET-SOFAST NMR for fast detection of structural compactness and heterogeneity along polypeptide chains. *Magn Reson Chem* 44 Spec No:S177–84. (Cited p. 11 and 19.)

[147] Schanda, P., Forge, V., and Brutscher, B., **2007**. Protein Folding and unfolding studied at atomic resolution by fast two-dimensional NMR spectroscopy. *Proc Natl Acad Sci U S A* 104:11257–11262. (Cited p. 9, 10, 30, 37, 57 and 58.)

[148] Schanda, P., Kupce, E., and Brutscher, B., **2005**. SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J Biomol NMR* 33:199–211. (Cited p. 11, 15, 19, 28 and 32.)

[149] Schanda, P., Van Melckebeke, H., and Brutscher, B., **2006**. Speeding up three-dimensional protein NMR experiments to a few minutes. *J Am Chem Soc* 128:9042–9043. (Cited p. 61, 62, 63 and 73.)

[150] Schipper, D., **1995**. Structural studies of the teichoic acids from *Bacillus licheniformis*. *Carbohydr Res* 279:75–82. (Cited p. 148.)

[151] Schirner, K., Marles-Wright, J., Lewis, R. J., and Errington, J., **2009**. Distinct and essential morphogenic functions for wall- and lipo-teichoic acids in *Bacillus subtilis*. *EMBO* 28:830–842. (Cited p. 153.)

[152] Schmidt-Rohr, K. and Spiess, H.-W., **1994**. *Multidimensional Solid-State NMR and polymers*. Academic Press. (Cited p. 105.)

[153] Schulteherbruggen, T., Madi, Z. L., Sorensen, O. W., and Ernst, R. R., **1991**. Reduction of Multiplet Complexity in COSY-Type NMR-Spectra - the Bilinear and Planar COSY Experiments. *Mol Phys* 72:847–871. (Cited p. 68.)

[154] Sein, J., Giraud, N., Blackledge, M., and Emsley, L., **2007**. The role of (15)N CSA and CSA/dipole cross-correlation in (15)N relaxation in solid proteins. *J Magn Reson* 186:26–33. (Cited p. 113 and 117.)

[155] Selenko, P., Frueh, D. P., Elsaesser, S. J., Hass, W., Gygi, S. P., and Wagner, G., **2008**. In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. *Nat Struct Mol Biol* 15:321–329. (Cited p. 10.)

[156] Selenko, P. and Wagner, G., **2007**. Looking into live cells with in-cell NMR spectroscopy. *J Struct Biol* 158:244–253. (Cited p. 10.)

[157] Severin, A. and Tomasz, A., **1996**. Naturally occuring peptidoglycan variants of *Streptococcus pneumoniae*. *J Bacteriol* 178:168–174. (Cited p. 161.)

[158] Sharif, S., Singh, M., Kim, S. J., and Schaefer, J., **2008**. *Staphylococcus aureus* Peptidoglycan Tertiary Structure from Carbon-13 Spin Diffusion. *J Am Chem Soc* 131:7023–30. (Cited p. 123 and 126.)

[159] Shrot, Y. and Frydman, L., **2003**. Single scan NMR at arbitrary dimensions. *J Am Chem Soc* 125:11385–11396. (Cited p. 15 and 41.)

[160] Smith, M., Hu, H., and Shaka, A., **2001**. Improved broadband inversion performance for NMR in liquids. *J Magn Reson* 151:269–283. (Cited p. 49.)

[161] Sorensen, O. W., Eich, G. W., Levitt, M. H., Bodenhausen, G., and Ernst, R. R., **1983**. Product Operator-Fomalism for the Description of NMR Pulse Experiments. *Prog Nucl Magn Reson Spectrosc* 16:163–192. (Cited p. 62.)

[162] Spegt, P., Meurer, B., and Weill, G., **1985**. Deuterium resonance in the solid state. Experimental determination of moments from quadrupolar solid echo with corrections for finite pulse width and bandwidth. *Journal of Physics E: Scientific Instruments* 18:869–874. (Cited p. 114.)

[163] Sprangers, R., Velyvis, A., and Kay, L. E., **2007**. Solution NMR of supramolecular complexes: providing new insights into function. *Nat Methods* 4:697–703. (Cited p. 105.)

[164] Stearn, A. and Stearn, W. E., **1930**. The nature of the Gram compound and its bearing of the mechanism of staining. *J Bacteriol* 20:287–295. (Cited p. 86.)

[165] Szeverenyi, N., Sullivan, M., and Maciel, G., **1982**. Observation of spin exchange by two-dimensional fourier transform ^{13}C cross polarization-magic-angle spinning. *J Magn Reson* 47:462–475. (Cited p. 118, 121 and 132.)

[166] Szyperski, T., Wider, G., Bushweller, J. H., and Wüthrich, K., **1993**. 3D 13C-15N-heteronuclear two-spin coherence spectroscopy for polypeptide backbone assignments in 13C-15N-double-labeled proteins. *J Biomol NMR* 3:127–32. (Cited p. 15.)

[167] Tollinger, M., Skrynnikov, N. R., Mulder, F. A., Forman-Kay, J. D., and Kay, L. E., **2001**. Slow dynamics in folded and unfolded states of an SH3 domain. *J Am Chem Soc* 123:11341–52. (Cited p. 75.)

[168] Tycko, R., **1994**. *Nuclear Magnetic Resonance Probes of Molecular Dynamics*. Kluwer Academic Publishers. (Cited p. 113 and 117.)

[169] Van Melckebeke, H., Simorre, J.-P., and Brutscher, B., **2004**. Amino acid-type edited NMR experiments for methyl-methyl distance measurement in 13C-labeled proteins. *J Am Chem Soc* 126:9584–91. (Cited p. 15, 23 and 24.)

[170] Van-Nuland, N., Forge, V., Balbach, J., and Dobson, C., **1998**. Real-time NMR studies of protein folding. *Acc Chem Res* 31:773–780. (Cited p. 9 and 27.)

[171] Vermeer, L. S., de Groot, B. L., Réat, V., Milon, A., and Czaplicki, J., **2007**. Acyl chain order parameter profiles in phospholipid bilayers: computation from molecular dynamics simulations and comparison with 2H NMR experiments. *Eur Biophys J* 36:919–31. (Cited p. 109.)

[172] Veshtort, M. and Griffin, R. G., **2004**. High-performance selective excitation pulses for solid- and liquid-state NMR spectroscopy. *Chemphyschem* 5:834–50. (Cited p. 64, 66 and 67.)

[173] Vollmer, W. and Bertsche, U., **2008**. Murein (peptidoglycan) structure, architecture and biosynthesis in Escherichia coli. *Biochim Biophys Acta* 1778:1714–1734. (Cited p. 87, 89 and 132.)

[174] Vollmer, W., Blanot, D., and de Pedro, M., **2008**. Peptidoglycan structure and architecture. *FEMS Microbiol Rev* 32:149–167. (Cited p. 86, 87, 89, 91 and 92.)

[175] Vollmer, W. and Hoeltje, J.-V., **2004**. The architecture of the murein (peptidoglycan) in gram-negative bacteria: vertical scaffold or horizontal layer(s)? *J Bacteriol* 186:5978–87. (Cited p. 87, 94 and 97.)

[176] Wang, C., Grey, M. J., and Palmer, A. G. r., **2001**. CPMG sequences with enhanced sensitivity to chemical exchange. *J Biomol NMR* 21:361–366. (Cited p. 76 and 77.)

[177] Warth, A. D. and Strominger, J. L., **1971**. Structure of the peptidoglycan from vegetative cell walls of *Bacillus subtilis*. *Biochemistry* 10:4349–4358. (Cited p. 89 and 134.)

[178] Weidel, W. and Pelzer, H., **1964**. Bagshaped-macromolecules a new outlook on bacterial cell walls. *Adv Enzymol Relat Areas Mol Biol* 26:193–232. (Cited p. 86.)

[179] Weidenmaier, C., Kokai-Kun, J. F., Kristian, S. A., Chanturiya, T., Kalbacher, H., Gross, M., Nicholson, G., Neumeister, B., Mond, J. J., and Peschel, A., **2004**. Role of teichoic acids in *Staphylococcus aureus* nasal colonization, a major risk factor in nosocomial infections. *Nat Med* 10:243–245. (Cited p. 99.)

[180] Weidenmaier, C. and Peschel, A., **2008**. Teichoic acids and related cell-wall glycopolymers in Gram-positive physiology and host interactions. *Nat Rev Micro* 6:276–287. 10.1038/nrmicro1861. (Cited p. 99 and 154.)

[181] Wickham, J. R., Jeffrey, L., Haley, L., Kashtanov, S., Khandogin, S., and Rice, C. V., **2009**. Revisiting magnesium chelation by teichoic acid with phosphorus solid-state NMR and theoretical calculations. *J Phys Chem B* 113:2177–2183. (Cited p. 101.)

[182] Wickham, J. R. and Rice, C. V., **2008**. Solid-State NMR studies of bacterial lipoteichoic acid adsorption on different surfaces. *Solid State Nucl Magn Reson* 34:154–161. (Cited p. 101, 114 and 152.)

[183] Zeeb, M. and Balbach, J., **2004**. Protein folding studied by real-time NMR spectroscopy. *Methods* 34:65–74. (Cited p. 9 and 27.)

ABSTRACT

Standard multidimensional NMR methods required to yield sufficient atom-resolution are time-consuming rendering real-time analysis difficult. The first part of this thesis treats recent advances in the field of fast NMR. The first part of this thesis is concerned with the optimization and extension of longitudinal relaxation enhanced pulse sequences in the context of SOFAST and BEST type experiments. The second part of this thesis deals with the bacterial cell wall of Gram-negative and Gram-positive bacteria. Because of its size and non cristalline character we applied solid state NMR to study it. The high-quality solid-state NMR spectra allow atom-resolved investigation of structure and dynamics of the peptidoglycan and its covalentely bound teichoic acids (WTA). The determination of dynamical properties of the peptidoglycan can also be used to study peptidoglycan-protein interactions and complexation with divalent ions.

RÉSUMÉ

Les méthodes RMN multidimensionnelles requises pour l'obtention d'une résolution atomique élevée sont relativement couteuses en temps expérimental, ce qui complique considérablement leur application à l'analyse d'échantillons en temps réel. La première partie de cette thèse traite des progrès récents dans le domaine de l'acquisition rapide des spectres de RMN . La deuxième partie concerne la paroi cellulaire des bactéries Gram-négatives et Gram-positives. En raison de son poids moléculaire élevé et de son caractère non cristallin, nous avons appliqué la RMN du solide pour l'étudier. La qualité exceptionnelle des spectres de RMN solide permet l'étude, à résolution atomique, de la structure et de la dynamique du peptidoglycane et des acides téchoïques qui se lient de manière covalente au peptidoglycan (WTA). La détermination des propriétés dynamiques du peptidoglycane est aussi utilisée pour étudier les interactions entre protéines et peptidoglycane et la complexation avec des ions divalents.

Mots-Clés: RMN, biologie structurale, paroi bactérienne, acides téchoïques, RMN du solid, RMN, rapide, SOFAST, BEST, relaxation dispersion
Laboratoire de thèse: Institut de Biologie Structurale Jean-Pierre Ebel, UMR 5075, CEA.; CNRS.; UJF. Laboratoire de Résonance Magnétique, 41, avenue Jules Horowitz - 38027 GRENOBLE CEDEX 1